

TCTGTGGTTT GTCAGACCTG CAAGCAAGCC CCCTGCTGGG GAAGCCTAGG TGTCTTGAG 2280  
 CTGAACCGCA CTGAAGAACT CTTGTCTCA CTGGCTGATG CAGCAGAACT CTTGGGAAAT 2340  
 GTCTTAGTCC TGCAGAACTA GGAGTCACCA GATGATGCAG AGTTGAGATC ATCATTGCAA 2400  
 AGTTCTCTGT TCCTGAGGAA CTAAATTTAA GGAAAAATG GGATTTTGT TTAGAGTTGG 2460  
 AAAAAAGCC TGATTAAAGA GTTTCTGCCT GTTAAAAAA AAAAAAAA AAAAA

Seq ID NO: 97 Protein sequence:  
 Protein Accession #: NP\_542399.1

1 11 21 31 41 51  
 | | | | |  
 MSGRRTRSGG AAQRSGPRAP SPTKPLRRSQ RKSGSELPSI LPEIWPKTPS AA AVRKP IVL 60  
 KRIVAHAVEV PAVQSPRRSP RISFFLEKEN EPPGRELTKE DLFKTHSVPA TPTSTPVPNP 120  
 EAESSSKEGE LDARDLEMSK KVRRSYSRLE TLGSASTSTP GRRSCFGFEG LLGAEDLSGV 180  
 SPVVCCKLTE VPRVCAPKPA PDMTLPGISP PPEKQKRKKK KMPEILKTEL DEWAAAMNAE 240  
 FEAAEQFDLL VE

Seq ID NO: 98 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 58-12444

1 11 21 31 41 51  
 | | | | |  
 GGGGCATTTC CGGGTCCGGG CCGAGCGGGC GCACGCGCGG GAGCGGGACT CGGCGGCATG 60  
 GCGGGCTCCG GAGCCGGTGT GCGTTGCTCC CTGCTCGCGC TGCAGGAGAC CTTGTCCGCT 120  
 GCGGACCGCT GCGGTGCTGC CCTGGCCGGT CATCAACTGA TCCGCGGCCT GGGGCAGGAA 180  
 TCGTCTCTGA GCAGCAGCCC CGCGGTGCTG GCATTACAGA CATCTTTAGT TTTTCCAGA 240  
 GATTTTCGGT TGCTTGTAAT TGTCCGGAAG TCACTCAACA GTATTGAATT TCGTGAATGT 300  
 AGAGAAGAAA TCCTAAAGTT TTTATGTATT TTCTAGAAA AAATGGGCCA GAAGATCGCA 360  
 CCTTACTCTG TTGAAATTAA GAACACTTGT ACCAGTGT TT ATACAAAAGA TAGAGCTGCT 420  
 AAATGTAAAA TTCCAGCCCT GGACCTTCTT ATTAAGTTAC TTCAGACTTT TAGAAGTTCT 480  
 AGACTCATGG ATGAATTTAA AATTGGAGAA TTATTTAGTA AATTCTATGG AGAACTTGCA 540  
 TTGAAAAAAA AAATACCAGA TACAGTTTTA GAAAAAGTAT ATGAGCTCCT AGGATTATTG 600  
 GGTGAAGTTC ATCCTAGTGA GATGATAAAT AATGCAGAAA ACCTGTTCCG CGCTTTTCTG 660  
 GGTGAAGTTA AGACCCAGAT GACATCAGCA GTAAGAGAGC CCAAACTACC TGTCTGGCA 720  
 GGATGTCTGA AGGGGTGTC CTCACTTCTG TGCAACTTCA CTAAGTCCAT GGAAGAAGAT 780  
 CCCCAGACTT CAAGGGAGAT TTTTAATTTT GTACTAAAG CAATTCGTCC TCAGATTGAT 840  
 CTGAAGAGAT ATGCTGTGCC CTCAGCTGGC TTGCGCCTAT TTGCCCTGCA TGCATCTCAG 900  
 TTTAGCACCT GCCTTCTGGA CAACTACGTG TCTCTATTG AAGTCTTGT AAAGTGGTGT 960  
 GCCCACACAA ATGTAGAATT GAAAAAGCT GCACTTTCAG CCTGGAATC CTTTCTGAA 1020  
 CAGGTTCTA ATATGGTGGC GAAAAATGCA GAAATGCATA AAAATAAACT GCAGTACTTT 1080  
 ATGGAGCAGT TTTATGGAAT CATCAGAAAT GTGGATTGCA ACAACAAGGA GTTATCTATT 1140  
 GCTATCCGTG GATATGGACT TTTTGCAGGA CCGTGCAAGG TTATAAACGC AAAAGATGTT 1200  
 GACTTCATGT ACGTTGAGCT CATTACAGCGC TGCAAGCAGA TGTTCCTCAC CCAGACAGAC 1260  
 ACTGGTGACG ACCGTGTTTA TCAGATGCCA AGCTTCCTCC AGTCTGTTGC AAGCGTCTTG 1320  
 CTGTACCTTG ACACAGTTCC TGAGGTGTAT ACTCCAGTTC TGGAGCACCT CGTGGTGATG 1380  
 CAGATAGACA GTTTCCACA GTACAGTCCA AAAATGCAGC TGGTGTGTTG CAGAGCCATA 1440  
 GTGAAGGTGT TCCTAGCTTT GGCAGCAAAA GGGCCAGTTC TCAGGAATTG CATTAGTACT 1500  
 GTGGTGATC AGGGTTTAA CAGAATATGT TCTAAACCAG TGGTCTTCC AAAGGGCCCT 1560  
 GAGTCTGAAT CTGAAGACCA CCGTGCTTCA GGGGAAGTCA GAACTGGCAA ATGGAAGGTG 1620  
 CCCACATACA AAGACTACGT GGATCTCTTC AGACATCTCC TGAGCTCTGA CCAGATGATG 1680  
 GATTCTATTT TAGCAGATGA AGCATTTTTT TCTGTGAATT CCTCCAGTGA AAGTCTGAAT 1740  
 CATTCTATTT ATGATGAATT TGTAAATCC TGTGTGAAGA TTGTTGAGAA ATTGGATCTT 1800  
 ACACTTGA AAA TACAGACTGT TGGGGAACAA GAGAATGGAG ATGAGGCGCC TGGTGTGTTG 1860  
 ATGATCCCAA CTTAGATCC AGCGGCTAAC TTGCATCCAG CTAAACCTAA AGATTTTTCG 1920  
 GCTTTCATTA ACCTGGTGGA ATTTTGCAGA GAGATTCTCC CTGAGAAACA AGCAGAAATT 1980  
 TTTGAACCAT GGGTGTACTC ATTTTCATAT GAATTAATTT TGCAATCTAC AAGGTTGCC 2040  
 CTCATCAGTG GTTCTACAA ATTGCTTCT ATTACAGTAA GAAATGCCAA GAAAATAAAA 2100  
 TATTTGAGG GAGTAGTCC AAAGAACTCT AAACACTCTC CTGAAGACCC AGAAAAGTAT 2160  
 TCTTGCTTTG CTTTATTTGT GAAATTTGGC AAAGAGGTGG CAGTTAAAT GAAGCAGTAC 2220  
 AAAGATGAAC TTTTGGCCTC TTGTTTGACC TTTCTTCTGT CCTTGCCACA CAACATCATT 2280  
 GAACTCGATG TTAGAGCCTA CGTTCCTGCA CTGCAGATGG CTTTCAAAC GGGCCTGAGC 2340  
 TATACCCCTT TGGCAGAAGT AGGCCTGAAT GCTCTAGAAG AATGGTCAAT TTATATTGAC 2400  
 AGACATGTAA TGCAGCCTTA TTACAAAGAC ATTCTCCCCT GCCTGGATGG ATACCTGAAG 2460  
 ACTTCAGCCT TGTCAGATGA GACCAAGAAG AACTGGGAAG TGTGAGCTCT TTCTCGGGCT 2520  
 GCCCAGAAAAG GATTTAATAA AGTGGTGTGA AAGCATCTGA AGAAGACAAA GAACCTTTCA 2580  
 TCAAACGAAG CAATATCCTT AGAAGAAATA AGAATTAGAG TAGTACAAAT GCTTGGATCT 2640  
 CTAGGAGGAC AAATAACAA AAATCTTCTG ACAGTCACGT CCTCAGATGA GATGATGAAG 2700  
 AGCTATGTGG CCTGGGACAG AGAGAAGCGG CTGAGCTTTG CAGTGCCCTT TAGAGAGATG 2760  
 AAACCTGTCA TTTTCTGGA TGTGTTCTG CCTCGAGTCA CAGAATTAGC GCTCACAGCC 2820  
 AGTGACAGAC AAATAAAGT TGCAGCCTGT GAACTTTTAC ATAGCATGGT TATGTTTATG 2880  
 TTGGGCAAAG CCACGCAGAT GCCAGAAGGG GGACAGGGAG CCCCACCCAT GTACCAGCTC 2940  
 TATAAGCGGA CGTTTCTGT GCTGCTTCGA CTTGCGTGTG ATGTTGATCA GGTGACAAGG 3000  
 CAACTGTATG AGCCACTAGT TATGCAGCTG ATTCAGTGGT TCACTAACA CAAGAAATT 3060  
 GAAAGTCAGG ATACTGTTGC CTTACTAGAA GCTATATTGG ATGGAATTGT GGACCTGTT 3120  
 GACAGTACTT TAAGAGATTT TTGTGGTCCG TGTATTCGAG AATTCCTTAA ATGGTCCATT 3180  
 AAGCAAATAA CACCACAGCA GCAGGAGAAG AGTCCAGTAA ACACCAAATC GCTTTTCAAG 3240  
 CGACTTTATA GCCTTGCGCT TCACCCCAAT GCTTTCAAGA GGCTGGGAGC ATCACTTGCC 3300  
 TTAATAATA TCTACAGGGA ATTCAGGGAA GAAGAGTCTC TGGTGAACA GTTTGTGTTT 3360  
 GAAGCCTTGG TGATATACAT GGAGAGTCTG GCCTTAGCAC ATGCAGATGA GAAGTCTTA 3420  
 GGTACAATTC AACAGTGTG TGATGCCATT GATCACCTAT GCCGCATCAT TGAAAAGAAG 3480  
 CATGTTTCTT TAAATAAAGC AAAGAAACGA CGTTTGCCGC GAGGATTTCC ACCTTCCGCA 3540  
 TCATTGTGTT TATTGGATCT GTCAAGTGG CTTTTAGCTC ATTGTGGGAG GCCCAGACA 3600  
 GAATTCGAC ACAATCCAT TGAACCTTT TTAATAATCG TTCCTTTATT GCCAGGCAAC 3660  
 AGATCCCTA ATTTGTGGCT GAAAGATGTT CTCAAGGAAG AAGGTGTCTC TTTTCTCATC 3720  
 AACACCTTTG AGGGGGGTGG CTGTGGCCAG CCCTCGGGCA TCCTGGCCCA GCCCACCCTC 3780  
 TTGTACCTTC GGGGGCCATT CAGCCTGCAG GCCACGCTAT GCTGGCTGGA CCTGCTCCTG 3840



	GCCGCGTTGG	AGTGCTACAA	CACGTTTCATT	GGCGAGAGAA	CTGTAGGAGC	GCTCCAGGTC	3900
	CTAGGTACTG	AAGCCCAGTC	TTCACTTTTG	AAAGCAGTGG	CTTTCCTTCT	AGAAAGCATT	3960
	GCCATGCATG	ACATTATAGC	AGCAGAAAAG	TGCTTTGGCA	CTGGGGCAGC	AGGTAACAGA	4020
5	ACAAGCCCAC	AAGAGGGAGA	AAGGTACAAC	TACAGCAAAT	GCACCGTTGT	GGTCCGGATT	4080
	ATGGAGTTTA	CCACGACTCT	GCTAAACACC	TCCCCGGAAG	GATGGAAGCT	CCTGAAGAAG	4140
	GACTTGTGTA	ATACACACCT	GATGAGAGTC	CTGGTGCAGA	CGCTGTGTGA	GCCCCGAAGC	4200
	ATAGGTTTCA	ACATCGGAGA	CGTCCAGGTT	ATGGCTCATC	TTCCTGATGT	TTGTGTGAAT	4260
	CTGATGAAAG	CTCTAAAGAT	GTCCCCATAC	AAAGATATCC	TAGAGACCCA	TCTGAGAGAG	4320
10	AAAATAACAG	CACAGAGCAT	TGAGGAGCTT	TGTGCCGTCA	ACTTGTATGG	CCCTGACGCG	4380
	CAAGTGGACA	GGAGCAGGCT	GGCTGCTGTT	GTGTCTGCCT	GTAAACAGCT	TCACAGAGCT	4440
	GGGCTTCTGC	ATAATATATT	ACCGTCTCAG	TCCACAGATT	TGCATCATT	TGTTGGCACA	4500
	GAACTTCTTT	CCCTGGTTTA	TAAAGGCATT	GCCCCGAG	ATGAGAGACA	GTGTCTGCCT	4560
	TCTCTAGACC	TCAGTTGTAA	GCAGCTGGCC	AGCGGACTTC	TGGAGTTAGC	CTTTGCTTTT	4620
	GGAGGACTGT	GTGAGCGCCT	TGTGAGTCTT	CTCCTGAACC	CAGCGGTGCT	GTCCACGGCG	4680
15	TCCTTGGGCA	GCTCACAGGG	CAGCGTCATC	CAGTTCTCCC	ATGGGGAGTA	TTTCTATAGC	4740
	TTGTTCTCAG	AAACGATCAA	CACGGAATTA	TTGAAAAATC	TGGATCTTGC	TGTATTGGAG	4800
	CTCATGCACT	CTTCAGTGGA	TAATACCAAA	ATGGTGAGTG	CCGTTTTGAA	CGGCATGTTA	4860
	GACCAGAGCT	TCAGGGAGCG	AGCAAACCAG	AAACACCAAG	GACTGAAACT	TGCGACTACA	4920
20	ATTCTGCAAC	ACTGGAAGAA	GTGTGATTCA	TGGTGGGCCA	AAGATTCCCC	TCTCGAAACT	4980
	AAAATGGCAG	TGCTGGCCTT	ACTGGCAAAA	ATTTTACAGA	TTGATTTCATC	TGTATCTTTT	5040
	AATACAAGTC	ATGGTTTCATT	CCCTGAAGTC	TTTACAACAT	ATATTAGTCT	ACTTGCTGAC	5100
	ACAAAGCTGG	ATCTACATTT	AAAGGGCCAA	GCTGTCACTC	TTCTTCCATT	CTTACCAGC	5160
	CTCACTGGAG	GCAGTCTGGA	GGAACCTTAGA	CGTGTCTGCG	AGCAGCTCAT	CGTTGCTCAC	5220
25	TTCCCCATGC	AGTCCAGGGA	ATTTCTCTCA	GGAACCTCCG	GGTTCAATAA	TTATGTGGAC	5280
	TGCATGAAAA	AGTTTCTAGA	TGCATTGGAA	TTATCTCAAA	GCCCTATGTT	GTTGGAATTG	5340
	ATGACAGAAG	TTCTTTGTCT	GGAACAGCAG	CATGTCATGG	AAGAATTATT	TCAATCCAGT	5400
	TTCAGGAGGA	TTGCCAGAAG	GGGTTTCATG	GTCAACAAAG	TAGGCCCTTCT	GGAAAGCGTG	5460
	TATGAAATGT	TCAGGAAGGA	TGACCCCGCG	CTAAGTTTCA	CACGCCAGTC	CTTTGTGGAC	5520
30	CGCTCCCTCC	TCACCTGTCT	GTGGCACTGT	AGCCTGGATG	CTTTGAGAGA	ATTCTTCAGC	5580
	ACAATTGTGG	TGGATGCCAT	TGATGTGTTG	AAGTCCAGGT	TTACAAAGCT	AAATGAATCT	5640
	ACCTTTGATA	CTCAAATCAC	CAAGAAGATG	GGCTACTATA	AGATTCTAGA	CGTGATGTAT	5700
	TCTCGCCTTC	CCAAAGATGA	TGTTTCATGCT	AAGGAATCAA	AAATTAATCA	AGTTTTCCAT	5760
	GGCTCGTGTA	TTACAGAAGG	AAATGAACCT	ACAAAGACAT	TGATTAAATT	GTGCTACGAT	5820
35	GCATTTACAG	AGAACATGGC	AGGATGAGAT	CAGCTGCTGG	AGAGGAGAAG	ACTTTACCAT	5880
	TGTGCAAGCAT	ACAACCTGCGC	CATATCTGTC	ATCTGCTGTG	TCTTCAATGA	GTTAAAATTT	5940
	TACCAAGGTT	TTCTGTTTAG	TGAAAAACCA	GAAAAGAACT	TGCTTATTTT	TGAAAATCTG	6000
	ATCGACCTGA	AGCGCCGCTA	TAATTTTCTT	GTAGAAAGTT	AGGTTCCCTAT	GGAAAGAAAG	6060
	AAAAAGTACA	TTGAAATTAG	GAAAGAAGCC	AGAGAAGCAG	CAAATGGGGA	TTGAGATGGT	6120
40	CCTTCTCTATA	TGCTTCTCCT	GTCTATTTTG	GCAGACAGTA	CCCTGAGTGA	GGAAATGAGT	6180
	CAATTTGATT	TCTCAACCGG	AGTTTCAGAG	TATTCATACA	GCTCCCAAGA	CCCTAGACCT	6240
	GCCACTGGTC	GTTTTGCGAG	ACGGGAGCAG	CGGGACCCCA	CGGTGCATGA	TGATGTGCTG	6300
	GAGCTGGAGA	TGGACGAGCT	CAATCGGCAT	GAGTGCATGG	CGCCCTGAC	GGCCCTGGTC	6360
	AAGCACATGC	ACAGAAGCCT	GGGCCCCGCT	CAAGGAGAAG	AGGATTCAGT	GCCAAGAGAT	6420
45	CTTCCTTCTT	GGATGAAATT	CCTCCATGGC	AAACTGGGAA	ATCCAATAGT	ACCATTAAAT	6480
	ATCCGTCTCT	TCTTAGCCAA	GCTTGTTATT	AATACAGAAG	AGGTCTTTCT	CCCTTACGCG	6540
	AAGCACTGGC	TGAGCCCTT	GCTGCAGCTG	GCTGCTTCTG	AAAACAATGG	AGGAGAAGGA	6600
	ATTCACCTACA	TGGTGGTTGA	GATAGTGCC	ACTATTCTTT	CATGGACAGG	CTTGGCCACT	6660
	CCAACAGGGG	TCCCTAAAGA	TGAAGTGTTA	GCAAATCGAT	TGCTTAATTT	CCTAATGAAA	6720
50	CATGTCTTTC	ATCCAAAAAG	AGCTGTGTTT	AGACACAACC	TTGAAATTAT	AAAGACCCTT	6780
	GTCGAGTGCT	GGAAGGATTG	TTTATCCATC	CCTTATAGGT	TAATATTTGA	AAAGTTTCTC	6840
	GGTAAAGATC	CTAATTCTAA	AGACAACCTCA	GTAGGGATTC	AATTGCTAGG	CATCGTGATG	6900
	GCCAAATGACC	TGCCTCCCTA	TGACCCACAG	TGTGGCATCC	AGAGTAGCGA	ATACTTCCAG	6960
	GCTTTGGTGA	ATAATATGTC	CTTTGTAAGA	TATAAAGAAG	TGTATGCCGC	TGCAGCAGAA	7020
55	GTTCTAGGAC	TTATACTTCG	ATATGTTATG	GAGAGAAAAA	ACATACTGGA	GGAGTCTCTG	7080
	TGTGAAGTGG	TTGCGAAACA	ATTGAAGCAA	CATCAGAATA	CTATGGAGGA	CAAGTTTATT	7140
	GTGTGCTTGA	ACAAAGTGAC	CAAGAGCTTC	CCTCCTCTTG	CAGACAGGTT	CATGAATGCT	7200
	GTGTTCTTTC	TGCTGCCAAA	ATTTTCATGGA	GTGTTGAAAA	CACTCTGTCT	GGAGGTGGTA	7260
	CTTTGTCTGT	TGGAGGGAAT	GACAGAGCTG	TACTTCCAGT	TAAAGAGCAA	GGACTTCTGT	7320
60	CAAGTCATGA	GACATAGAGA	TGATGAAAGA	CAAAAAGTAT	GTTTGGACAT	AATTTATAAG	7380
	ATGATGCCAA	AGTTAAAACC	AGTAGAATC	CGAGAACTTC	TGAACCCCGT	TGTGGAATTC	7440
	GTTTCCCATC	CTTCTACAAC	ATGTAGGGAA	CAAATGTATA	ATATTCTCAT	GTGGATTCAAT	7500
	GATAATTACA	GAGATCCAGA	AAGTGAGACA	GATAATGACT	CCCAGGAAAT	ATTTAAGTTG	7560
	GCAAAAGATG	TGCTGATTCA	AGGATTGATC	GATGAGAACC	CTGGACTTCA	ATTAATTATT	7620
65	CGAAATTTCT	GGAGCCATGA	AACTAGGTTA	CCTTCAAATA	CCTTGGACCG	GTTGCTGGCA	7680
	CTAAATTCCT	TATATTCTCC	TAAGATAGAA	GTGCACTTTT	TAAGTTTAGC	AACAAATTTT	7740
	CTGCTCGAAA	TGACCCAGAT	GAGCCAGAT	TATCCAAACC	CCATGTTCTG	GCATCCTCTG	7800
	TCAGAATGCG	AATTTTCAGGA	ATATAACCAT	GATTCTGATT	GGCGTTTCCG	AAGTACTGTT	7860
	CTCACTCCGA	TGTTTGTGGA	GACCCAGGCC	TCCCAGGGCA	CTCTCCAGAC	CCGTACCCAG	7920
70	GAAGGGTCCC	TCTCAGCTCG	CTGGCCAGTG	GCAGGGCAGA	TAAGGGCCAC	CCAGCAGCAG	7980
	CATGACTTCA	CACTGACACA	GACTGCAGAT	GGAAGAAGCT	CATTTGATTG	GCTGACCGGG	8040
	AGCAGCACTG	ACCCGCTGGT	CGACCACACC	AGTCCCTCAT	CTGACTCCTT	GCTGTTTGCC	8100
	CACAAGAGGA	GTGAAAGGTT	ACAGAGAGCA	CCCTTGAAGT	CAGTGGGGCC	TGATTTTGGG	8160
	AAAAAAGGC	TGGGCCTTCC	AGGGGACGAG	GTGGATAACA	AAGTGAAAGG	TGCGGCCGGC	8220
75	CGGACGGACC	TACTACGACT	GCGCAGACGG	TTTATGAGGG	ACCAGGAGAA	GCTCAGTTTG	8280
	ATGTATGCCA	GAAAAGGCGT	TGCTGAGCAA	AAACGAGAGA	AGGAAATCAA	GAGTGAGTTA	8340
	AAAATGAAGC	AGGATGCCCA	GGTCGTTCTG	TACAGAAGCT	ACCGGCACGG	AGACCTTCCT	8400
	GACATTGAGA	TCAAGCACAG	CAGCCTCATC	ACCCGTTTAC	AGGCCGTGGC	CCAGAGGGAC	8460
	CAAATAATTG	CAAAACAGCT	CTTTAGCAGC	TTGTTTTCTG	GAATTTTGAA	AGAGATGGAT	8520
80	AAATTTAAGA	CACGTGCTGA	AAAAAACAAAC	ATCACTCAAA	AGTTGCTTCA	AGACTTCAAT	8580
	CGTTTTCTTA	ATACCACCTT	CTCTTTCTTT	CCACCTTTTG	TCTCTGTAT	TCAGGACATT	8640
	AGCTGTCAGC	ACGCAGCCCT	GCTGAGCCTC	GACCCAGCGG	CTGTTAGCGC	TGTTTGCCCTG	8700
	GCCAGCCTAC	AGCAGCCCGT	GGGCATCCGC	CTGCTAGAGG	AGGCTCTGCT	CCGCCTGCTG	8760
	CCTGCTGAGC	TGCCTGCCAA	GCGAGTCCGT	GGGAAGGCCC	GCTTCCCTCC	TGATGTCCTC	8820
85	AGATGGGTGG	AGCTTGCTAA	GCTATATAGA	TCAATTGGAG	AATACGACGT	CCTCCGTGGG	8880
	ATTTTATACCA	GTGAGATAGG	AACAAAGCAA	ATCACTCAGA	GTGCATTATT	AGCAGAAGCC	8940
	AGAAGTGATT	ATTCTGAAGC	TGCTAAGCAG	TATGATGAGG	CTCTCAATAA	ACAAGACTGG	9000
	GTAGATGGTG	AGCCACAGAA	AGCCGAGAAG	GATTTTTGGG	AACTTGCATC	CCTTGACTGT	9060



	TACAACCACC	TTGCTGAGTG	GAAATCACTT	GAATACTGTT	CTACAGCCAG	TATAGACAGT	9120
	GAGAACCCCC	CAGACCTAAA	TAAAATCTGG	AGTGAACCAT	TTTATCAGGA	AACATATCTA	9180
	CCTTACATGA	TCCGCAGCAA	GCTGAAGCTG	CTGCTCCAGG	GAGAGGCTGA	CCAGTCCCTG	9240
5	CTGACATTTA	TTGACAAAGC	TATGCACGGG	GAGCTCCAGA	AGGCGATTCT	AGAGCTTCAT	9300
	TACAGTCAAG	AGCTGAGTCT	GCTTTACCTC	CTGCAAGATG	ATGTTGACAG	AGCCAAATAT	9360
	TACATTCAAA	ATGGCATTCA	GAGTTTTATG	CAGAATTATT	CTAGTATTGA	TGTCCTCTTA	9420
	CACCAAAGTA	GACTCACCAA	ATTGCAGTCT	GTACAGGCTT	TAACAGAAAT	TCAGGAGTTC	9480
	ATCAGCTTTA	TAAGCAAACA	AGGCAATTTA	TCATCTCAAG	TTCCCCTTAA	GAGACTTCTG	9540
10	AACACCTGGA	CAAACAGATA	TCCAGATGCT	AAAATGGACC	CAATGAACAT	CTGGGATGAC	9600
	ATCATCACAA	ATCGATGTTT	CTTTCTCAGC	AAAATAGAGG	AGAAGCTTAC	CCCTCTTCCA	9660
	GAAGATAATA	GTATGAATGT	GGATCAAGAT	GGAGACCCCA	GTGACAGGAT	GGAAGTGCAA	9720
	GAGCAGGAAG	AAGATATCAG	CTCCCTGATC	AGGAGTTGCA	AGTTTTCCAT	GAAAAATGAAG	9780
	ATGATAGACA	GTGCCCCGAA	GCAGAACAAT	TTCTCACTTG	CTATGAAACT	ACTGAAGGAG	9840
15	CTGCATAAAG	AGTCAAAAAC	CAGAGACGAT	TGGCTGGTGA	GCTGGGTGCA	GAGCTACTGC	9900
	CGCCTGAGCC	ACTGCCGGAG	CCGGTCCCAG	CGCTGCTCTG	AGCAGGTGCT	CACTGTGCTG	9960
	AAAACAGTCT	CTTTGTTGGA	TGAGAACAAC	GTGTCAAGCT	ACTTAAGCAA	AAATATTCTG	10020
	GCTTTCCGTG	ACCAGAACAT	TCTCTTGGGT	ACAACTTACA	GGATCATAGC	GAATGCTCTC	10080
	AGCAGTGAGC	CAGCCTGCCT	TGCTGAAATC	GAGGAGGACA	AGGCTAGAAG	AATCTTAGAG	10140
20	CTTTCTGGAT	CCAGTTTCTA	GGATTTCAGG	AAGGTGATCG	CGGGTCTGTA	CCAGAGAGCA	10200
	TTCCAGCACC	TCTCTGAGGC	TGTGCAGGCG	GCTGAGGAGG	AGGCCAGGCC	TCCCTCCTGG	10260
	AGCTGTGGGC	CTGCAGCTGG	GGTGATTGAT	GCTTACATGA	CGCTGGCAGA	TTTCTGTGAC	10320
	CAACAGCTGC	GCAAGGAGGA	AGAGAATGCA	TCAGTTATTG	ATTCTGCAGA	ACTGCAGGCG	10380
	TATCCAGCAC	TTGTGGTGGG	GAAAATGTTG	AAAGCTTTAA	AATTAAATTC	CAATGAAGCC	10440
25	AGATTGAAGT	TTCTTAGATT	ACTTCAGATT	ATAGAACGGT	ATCCAGAGGA	GACTTTGAGC	10500
	CTCATGACAA	AAGAGATCTC	TTCCGTTCCC	TGCTGGCAGT	TCATCAGCTG	GATCAGCCAC	10560
	ATGGTGGCCT	TACTGGACAA	AGACCAAGCC	GTTCGTGTTT	AGCACTCTGT	GGAAGAAATC	10620
	ACTGATAACT	ACCCGCAGGC	TATTGTTTTT	CCCTTCATCA	TAAGCAGCGA	AAGCTATTCC	10680
	TTCAAGGATA	CTTCTACTGG	TCATAAGAAT	AAGGAGTTTG	TGGCAAGGAT	TAAAAGTAAG	10740
30	TTGGATCAAG	GAGGAGTGAT	TCAAGATTTT	ATTAATGCCT	TAGATCAGCT	CTCTAATCCT	10800
	GAACGTCTCT	TTAAGGATTG	GAGCAATGAT	GTAAGAGCTG	AACTAGCAAA	AACCCCTGTA	10860
	AATAAAAAAA	ACATTGAAAA	AATGTATGAA	AGAATGTATG	CAGCCTTGGG	TGACCCAAAG	10920
	GCTCCAGGCC	TGGGGGCCTT	TAGAAGGAAG	TTTATTTCAGA	CTTTTGGAAA	AGAATTTGAT	10980
	AAACATTTTG	GGAAAGGAGG	TTCTAAACTA	CTGAGAATGA	AGCTCAGTGA	CTTCAACGAC	11040
35	ATTACCAACA	TGCTACTTTT	AAAAATGAAC	AAAGACTCAA	AGCCCCCTGG	GAATCTGAAA	11100
	GAATGTTTAC	CCTGGATGAG	CGACTTCAAA	GTGGAGTTCC	TGAGAAATGA	GCTGGAGATT	11160
	CCCGGTCAGT	ATGACGGTAG	GGGAAAGCCA	TTGCCAGAGT	ACCACGTGCG	AATCGCCGGG	11220
	TTTGATGAGC	GGGTGACAGT	CATGGCGTCT	CTGCCAAGGC	CCAAGCGCAT	CATCATCCGT	11280
	GGCCATGACG	AGAGGGAAACA	CCCTTTCTCT	GTGAAGGGTG	GCGAGGACCT	GCGGCAGGAC	11340
40	CAGCGCGTGG	AGCAGCTCTT	CCAGGTCTAT	AATGGGATCC	TGGCCCAAGA	CTCCGCTGCT	11400
	AGCCAGAGGG	CCCTGCAGCT	GAGGACCTAT	AGCGTTGTGC	CCATGACCTC	CAGGTTAGGA	11460
	TTAATTGAGT	GGCTTGAAAA	TACTGTTACC	TTGAAGGACC	TTCTTTTGAA	CACCATGTCC	11520
	CAAGAGGAGA	AGGCGGCTTA	CCTGAGTGAT	CCCAGGGCAC	CGCCGTGTGA	ATATAAAGAT	11580
	TGGCTGACAA	AAATGTCAGG	AAAACATGAT	GTTGGAGCTT	ACATGCTAAT	GTATAAGGGC	11640
45	GCTAATCGTA	CTGAAACAGT	CACGTCTTTT	AGAAAACGAG	AAAGTAAAGT	GCCTGCTGAT	11700
	CTCTTAAAGC	GGGCCTTCGT	GAGGATGAGT	ACAAGCCCTG	AGGCTTTCCT	GGCGCTCCGC	11760
	TCCCACTTCG	CCAGCTCTCA	CGCTCTGATA	TGCTCAGGCC	ACTGGATCCT	CGGGATTGGA	11820
	GACAGACATC	TGAACAACCT	TATGGTGGCC	ATGGAGACTG	GCGGCGTGAT	CGGGATCGAC	11880
	TTTGGGTCAT	CGTTTGGATC	CGCTACACAG	TTTCTGCCAG	TCCCTGAGTT	GATGCCTTTT	11940
50	CGGCTAACTC	GCCAGTTTAT	CAATCTGATG	TTACCAATGA	AAGAAACGGG	CCTTATGTAC	12000
	AGCATCATGG	TACACGCACT	CCGGGCCTTC	CGCTCAGACC	CTGGCCTGCT	CACCAACACC	12060
	ATGGATGTGT	TTGTCAAGGA	GCCCTCCTTT	GATTGGAAAA	ATTTTGAACA	GAAAATGCTG	12120
	AAAAAAGGAG	GGTCATGGAT	TCAAGAAATA	ACAGTTGCTG	AAAAAAATTG	GTACCCCGCA	12180
	CAGAAAATAT	GTTACGCTAA	GAGAAAGTTA	GCAGTGCCCA	ATCCAGCAGT	CATTACTTGT	12240
55	GATGAGCTAC	TCCTGGGTCA	TGAGAAGGCC	CCTGCCCTCA	GAGACTATGT	GGCTGTGGCA	12300
	CGAGGAAGCA	AAGATCACAA	CATTCTGTCC	CAAGAACCAG	AGAGTGGGCT	TTCAGAAGAG	12360
	ACTCAAGTGA	AGTGCCCTGAT	GGACCAGGCA	ACAGACCCCA	ACATCCTTGG	CAGAACCCTG	12420
	GAAGGATGGG	AGCCCTGGAT	GTGAGGTCTG	TGGGAGTCTG	CAGATAGAAA	GCATTACATT	12480
	GTTTAAAGAA	TCTACTATAC	TTTGGTTGGC	AGCATTCCTT	GAGCTGATTT	TCCTGAAACA	12540
60	CTAAAGAGAA	ATGTCTTTTG	TGCTACAGTT	TCGTAGCATG	AGTTTAAATC	AAGATTATGA	12600
	TGAGTAAATG	TGTATGGGTT	AAATCAAAGA	TAAGGTTATA	GTAACATCAA	AGATTAGGTG	12660
	AGGTTTATAG	AAAGATAGAT	ATCCAGGCTT	ACCAAAGTAT	TAAGTCAAGA	ATATAATATG	12720
	TGATCAGCTT	TCAAAGCATT	TACAAGTGCT	GCAAGTTAGT	GAAACAGCTG	TCTCCGTAAA	12780
	TGGAGGAAAT	GTGGGGGAAGC	CTTGGAATGC	CTTCTGCTT	CTGGCACATT	GGAAAGCACA	12840
65	CTCAGAAGGC	TTTCATACCA	AGATTTTGGG	AGAGTAAAGC	TAAGTATAGT	TGATGTAACA	12900
	TTGTAGAAGC	AGCATAGGAA	CAATAAGAAC	AATAGGTAAA	GCTATAATTA	TGGCTTATAT	12960
	TTAGAAATGA	CTGCATTTGA	TATTTTAGGA	TATTTTCTTA	GGTTTTTTCC	TTTCATTTTA	13020
	TTCTCTTCTA	GTTTTGACAT	TTTATGATAG	ATTTGCTCTC	TAGAAGGAAA	CGTCTTTATT	13080
	TAGGAGGGCA	AAAATTTTGG	TCATAGCATT	CACTTTTGTCT	ATTCCAATCT	ACAACCTGGA	13140
70	GATACATAAA	AGTGCTTTGC	ATTGAATTTG	GGATAACTTC	AAAAATCCCA	TGGTTGTTGT	13200
	TAGGGATAGT	ACTAAGCATT	TCAGTTCCAG	GAGAATAAAA	GAAATTCCTA	TTTGAATGA	13260
	ATTCCTCATT	TGGAGGAAAA	AAAGCATGCA	TTCTAGCACA	ACAAGATGAA	ATTATGGAAT	13320
	ACAAAAGTGG	CTCCTTCCCA	TGTGCAGTCC	CTGTCCCCCC	CCGCCAGTCC	TCCACACCCA	13380
	AACTGTTTCT	GATTGGCTTT	TAGCTTTTTG	TTGTTTTTTT	TTTTCTTCT	AACACTTGTA	13440
75	TTTGGAGGCT	CTTCTGTGAT	TTTGAGAAGT	ATACTCTTGA	GTGTTTAATA	AAGTTTTTTT	13500
	CCAAAAGTA						

Seq ID NO: 99 Protein sequence:  
Protein Accession #: NP\_008835.5

80	1	11	21	31	41	51	
	MAGSGAGVRC	SLLRLQETLS	AADRCGAALA	GHQLIRGLGQ	ECVLSSSPAV	LALQTSLVFS	60
	RDFGLLVFVR	KSLNSIEFRE	CREEILKFLC	IFLEKMGQKI	APYSVEIKNT	CTSVYTKDRA	120
	AKCKIPALDL	LIKLLQTFRS	SRLMDEFKIG	ELFSKFYGEI	ALKKKIPDTV	LEKVYELLGL	180
85	LGEVHPSEMI	NNAENLFRAP	LGEIKTQMTS	AVREPKLPLV	AGCLKGLSSL	LCNFTKSMEE	240
	DPQTSREIFN	FVLKAIRPQI	DLKRYAVPSA	GLRLFALHAS	QFSTCLLDNY	VSLFEVLLKW	300
	CAHTNVELKK	AALSALSFLL	KQVSNMVAKN	AEMHKNKLQY	FMEQFYGIIR	NVDSNNKELS	360



	IAIRGYGLFA	GPCKVINAKD	VDFMYVELIQ	RCKQMFLTQT	DTGDDRVYQM	PSFLQSVASV	420
	LLYLDTVPEV	YTFVLEHLVV	MQIDSFPQYS	PKMQLVCCRA	IVKVFLALAA	KGPVLRNCIS	480
	TVVHQGLIRI	CSKPVVLPKG	PESESEDHRA	SGEVRTGKWK	VPTYKDYVDL	FRHLLSSDQM	540
5	MDSILADEAF	FSVNSSSESL	NHLLYDEFVK	SVLKIVEKLD	LTLEIQTVEG	QENGDEAPGV	600
	WMIPTSDPAA	NLHPAKPKDF	SAFINLVEFC	REILPEKQAE	FFEPWVYSFS	YELILQSTRL	660
	PLISGFYKLL	SITVRNAKKI	KYFEGVSPKS	LKHSPEDEPK	YSCFALFVKF	GKEVAVKMKQ	720
	YKDELLASCL	TFLLSLPHNI	IELDVRAYVP	ALQMAFKLGL	SYTPLAEVGL	NALEEWISIYI	780
	DRHVMQPYK	DILPCLDGYL	KTSALSDETK	NNWEVSALSR	AAQKGFNKVV	LKHLKKTKNL	840
10	SSNEAISLEE	IRIRVVQMLG	SLGGQINKNL	LTVTSSDEMM	KSYVAWDREK	RLSFAVPFRE	900
	MKPVIFLDVF	LPRVTELALT	ASDRQTKVAA	CELLHSMVMF	MLGKATQMP	GGQGAPPMYQ	960
	LYKRTFPVLL	RLACDQVQVT	RQLYEPLVMQ	LIHWFTNNKK	FESQDITVALL	EAILDGIQVDP	1020
	VDSTLRDFCG	RCIREFLKWS	IKQITPQQQE	KSPVNTKSLF	KRLYSLALHP	NAFKRLGASL	1080
	AFNNIYREFR	EEESLVEQFV	FEALVIYMES	LALAHADEKS	LGTIQQCCDA	IDHLCRIIEK	1140
15	KHVSINKAKK	RRLPRGFPPS	ASLCLLDLVK	WLLAHCGRPQ	TECRHKSIEL	FYKFVPLLP	1200
	NRSPNLWLKD	VLKEEGVSFL	INTFEGGGCG	QPSGILAQPT	LLYLRGPFSL	QATLCWLDLL	1260
	LAALECYNTF	IGERTVGALQ	VLGTEAQSSL	LKAVAFFLES	IAMHDIIAAE	KCFGGAAGN	1320
	RTSPQEGERY	NYSKCTVVVR	IMEFTTTLLN	TSPEGWKLLK	KDLCNTHLMR	VLVQTLCEPA	1380
	SIGFNIGDVQ	VMAHLPDVCV	NLMKALKMSP	YKDILETHLR	EKITAQSIEE	LCAVNLYGPD	1440
20	AQVDRSLAA	VVSACKQLHR	AGLLHNILPS	QSTDHHSVVG	TELLSLVYKG	IAPGDERQCL	1500
	PSLDLSCQQL	ASGLLELAF	FGGLCERLVS	LLNPAVLST	ASLGSSQGSV	IHFSGHEYFY	1560
	SLFSETINTE	LLKNLDLAVL	ELMQSSVDNT	KMVSALVNGM	LDQSFRRERAN	QKHQGLKLAT	1620
	TILQHWKCD	SWWAKDSPL	TKMAVLALLA	KILQIDSSVS	FNTSHGSFPE	VFTTYISLLA	1680
	DTKLDLHLKG	QAVTLLPFFT	SLTGGSLLEL	RRVLEQLIVA	HFPMSREFP	PGTPRFNNYV	1740
25	DCMKKFLDAL	ELSQSPMLLE	LMTEVLCREQ	QHVMEELFQS	SFRRIARRGS	CVTQVGLLES	1800
	VYEMFRKDDP	RLSFTRQSFV	DRSLLTLLWH	CSLDALREFF	STIVVDAIDV	LKSRFTKLNE	1860
	STFDTQITTK	MGYYKILDVM	YSRLPKDDVH	AKESKINQVF	HGSCITEGNE	LTKTLIKLCY	1920
	DAFTENMAGE	NQLLERRRLY	HCAAYNCAIS	VICCVFNEK	FYQGFLEFSEK	PEKNLLIFEN	1980
	LIDLKRRYNF	PVEVEVPMER	KKKYIEIRKE	AREAANGSD	GPSYMSLSY	LADSTLSEEM	2040
30	SQDFDSTGVQ	SYSYSSQDPR	PATGRFRRRE	QRDPTVHDDV	LELEMDELNR	HECMAPLTAL	2100
	VKHMHRSLGP	PQGEEDSVPR	DLPSWMKFLH	GKLGNIPIVPL	NIRLFLAKLV	INTEEVFRPY	2160
	AKHWLSPLLQ	LAASENNGGE	GIHYMVVEIV	ATILSWTGLA	TPTGVPKDEV	LANRLNLFM	2220
	KHVFHPRKRAV	FRHNLEIIKT	LVECKWDCLS	IPYRLIFEKF	SGKDPNSKDN	SVGIQLLGIV	2280
	MANDLPPYDP	QCGIQSSEYF	QALVNNMSFV	RYKEYVAAAA	EVLGLILRYV	MERKNILEES	2340
35	LCELVAQQLK	QHONTMEDKF	IVCLNKVTKS	FPPLADRFMN	AVFFLLPKFH	GVLKTLCEV	2400
	VLCRVEGMT	LYFQLKSKDF	VQVMHRDDE	RQKVCLEIY	KMPKLPVE	LRELLNPVVE	2460
	FVSHPSTTCR	EQMYNIMWI	HDNYRDPSE	TDNDSQEIFK	LAKDVLIQGL	IDENPGLQLI	2520
	IRNFWSHETR	LPSNTLDRLL	ALNSLYSPKI	EVHFLSLATN	FLLEMTSMSP	DYPNPMFEHP	2580
	LSECEFOEYT	IDSDFRFRST	VLTPMFVETQ	ASQGTQTRT	QEGSLSARWP	VAGQIRATQQ	2640
40	QHDFTLTQTA	DGRSSFDWLT	GSSTDPLVDH	TSPSSDSLIF	AHKRSERLQR	APLKSVPDF	2700
	GKKRLGLPGD	EVDNKVKGA	GRDILLRLR	RFMRDQEKLS	LMYARKGVAE	QKREKEIKSE	2760
	LKMKQDAQVV	LYRSYRHGDL	PDIQIKHSSL	ITPLQAVQAR	DPIIAKQLFS	SLFSGILKEM	2820
	DKFKTLSEKN	NITQKLLQDF	NRFLNTTFSF	FPPFVSCIQD	ISCQHAALLS	LDPAAVSAGC	2880
	LASLQQPVGI	RLLEEALLRL	LPAELPAKRV	RGKARLPDVP	LRWVELAKLY	RSIGEYDVL	2940
45	GIFTSEIGTK	QITQSALLAE	ARSDYSEAAK	QYDEALNKQD	WVDGEPTAE	KDFWELASLD	3000
	CYNHLAEWKS	LEYCSTASID	SENPPDLNKI	WSEPFYQETY	LPYMIRSKLK	LLLQGEADQS	3060
	LLTFIDKAMH	GELQKAILLE	HYSQELSLLY	LLQDDVDRAK	YYIQNGIQSF	MQNYSSIDVL	3120
	LHQSLRLTKLQ	SVQALTEIQE	FISFISKQGN	LSSQVPLKRL	LNTWTNRYPD	AKMDPMNIWD	3180
	DIITNRCFFL	SKIEEKLTP	PEDNSMNVQ	DGDPDRMEV	QEQUEEISSL	IRSCFKSMKM	3240
50	KMIDSARKQN	NFSLAMKLLK	ELHKEKTRD	DWLVSQVSY	CRLSHCRSRS	QGCSEQVLT	3300
	LKTVSLLDEN	NVSSYLSKNI	LAFRDQNIL	GTTYRIIANA	LSSEPACLA	IEEDKARRIL	3360
	ELSGSSSEDS	EKVIAGLYQR	AFQHLSEAVQ	AAEEEAQPPS	WSCGPAAGVI	DAYMTLADFC	3420
	DQQLRKEEEN	ASVIDSAELQ	AYPALVVEKM	LKALKLNSNE	ARLKFPRLQ	IERYPEETL	3480
	SLMTKEISSV	PCWQFISWIS	HMVALDKDQ	AVAVQHSVEE	ITDNYPAIV	YPFIISSSEY	3540
55	SFKDTSIGHK	NKEFVARIKS	KLDQGGVIQD	FINALDQLSN	PELLFKDWSN	DVRAELAKTP	3600
	VNKNIEKMY	ERMYAALGDP	KAPGLGAFRR	KFIQTGKEF	DKHFGKGGSK	LLRMKLSDFN	3660
	DITNMLLLKM	NKDSKPPGNL	KECSPWMSDF	KVEFLRNELE	IPQYDGRGK	PLPEYHVRIA	3720
	GFDERVTMA	SLRRPKRIII	RGHDEREHPF	LVKGGEDLRQ	DQVEQLFQV	MNGILAQDSA	3780
	CSQALQLRT	YSVVPMTSRL	GLIEWLENTV	TLKDLLLNTM	SQEEKAAYLS	DPRAPPCEYK	3840
60	DWLTKMSGKH	DVGAYMLMYK	GARNRTEVTS	FRKRESKVPA	DLKRAFVRM	STSPEAFLLAL	3900
	RSHFASSHAL	ICISHWILGI	GDRHLNFMV	AMETGGVIGI	DFGHAFGSAT	QFLPVPELMP	3960
	FRLTRQFINL	MLPMKETGLM	YSIMVHALRA	FRSDPGLLTN	TMDVFVKEPS	FDWKNFEQKM	4020
	LKKGGSWIQE	INVAEKNWYP	RQKICYAKRK	LAGANPAVIT	CDELLLGHEK	APAFRDYVAV	4080
	ARGSKDHNIR	AQEPESGLSE	ETQVKCLMDQ	ATDPNILGRT	WEGWEPWM		

Seq ID NO: 100 DNA sequence  
Nucleic Acid Accession #: NM\_000673  
Coding sequence: 101-1225

70	1	11	21	31	41	51	
	ATGTGAAGGC	ACAAGCTGCT	GTTATATACA	ACAGAGTGAA	CTGAGCATCA	CTCAGAAAAA	60
	GTCTATGTTT	GCAGAAATAC	AGATCCAAGA	CAAAGACAGG	ATGGGCACTG	CTGGAAAAGT	120
	TATTAAATGC	AAAGCAGCTG	TGCTTTGGGA	GCAGAGCAA	CCCTTCTCCA	TTGAGGAAAT	180
75	AGAAGTTGCC	CCACCAAAGA	CTAAAGAAGT	TCGCATTAAG	ATTTTGGCCA	CAGGAATCTG	240
	TCGCACAGAT	GACCATGTGA	TAAAAGGAAC	AATGGTGTCC	AAGTTTCCAG	TGATTGTGGG	300
	ACATGAGGCA	ACTGGGATTG	TAGAGAGCAT	TGGAGAAGGA	GTGACTACAG	TGAAACCAGG	360
	TGACAAAGTC	ATCCCTCTCT	TTCTGCCACA	ATGTAGAGAA	TGCAATGCTT	GTGCAACCCC	420
	AGATGGCAAC	CTTTGCATTA	GGAGCATAT	TACTGGTCGT	GGAGTACTGG	CTGATGGCAC	480
80	CACCAAGTTT	ACATGCAAGG	GCAAACCATG	TACCCACTTC	ATGAACACCA	GTACATTTAC	540
	CGAGTACACA	GTGGTGGATG	AATCTTCTGT	TGCTAAGATT	GATGATGCAG	CTCCTCTGTA	600
	GAAAGTCTGT	TTAATTGGCT	GTGGGTTTTT	CACTGGATAT	GGCGCTGCTG	TTAAAACCTG	660
	CAAGGTCAAA	CCTGGTTCCA	CTTGCCTCGT	CTTTGGCCTG	GGAGGAGTTG	GCCTGTCTAGT	720
	CATCATGGGC	TGTAAGTCAG	CTGGTGCATC	TAGGATCATT	GGGATTGACC	TCAACAAAGA	780
	CAAATTTGAG	AAGGCCATGG	CTGTAGGTGC	CACGTAGTGT	ATCAGTCCCA	AGGACTCTAC	840
85	CAAACTTATC	AGTGAGGTGC	TGTGAGAAAT	GACAGGCAAC	AACGTGGGAT	ACACCTTTGA	900
	AGTTATTGGG	CATCTTGAAA	CCATGATTGA	TGCCCTGGCA	TCCTGCCACA	TGAACATATG	960
	GACCAGCGTG	GTTGTAGGAG	TTCCTCCATC	AGCCAAGATG	CTCACCTATG	ACCCGATGTT	1020



GCTCTTCACT GGACGCACAT GGAAGGGATG TGTCTTTGGA GGT TTGAAAA GCAGAGATGA 1080  
 TGTCCCAAAA CTAGTGACTG AGTTCCTGGC AAAGAAATTT GACCTGGACC AGTTGATAAC 1140  
 TCATGTTTTA CCATTTAAAA AAATCAGTGA AGGATTTGAG CTGCTCAATT CAGGACAAAG 1200  
 CATTGGAACG GTCCTGACGT TTTGAGATCC AAAGTGGCAG GAGGTCTGTG TTGTCATGGT 1260  
 5 GAACTGGAGT TTCTCTGTG AGAGTTCCTT CATCTGAAAT CATGTATCTG TCTCACAAT 1320  
 ACAAGCATAA GTAGAAGATT TGTGAAGAC ATAGAACCCT TATAAAGAAT TATTAACCTT 1380  
 TATAAACATT TAAAGTCTTG TGAGCACCTG GGAATTAGTA TAATAACAAT GTTAATATTT 1440  
 TTGATTTACA TTTTGTAAGG CTATAATTGT ATCTTTTAAG AAAACATACA CTTGGATTTC 1500  
 10 TATGTTGAAA TGGAGATTTT TAAGAGTTTT AACAGCTGC TGCAGATATA TAACTCAAAA 1560  
 CAGATATAGC GTATAAAGAT ATAGTAAATG CATCTCCAG AGTAATATTC ACTTAACACA 1620  
 TTGAACTAT TATTTT TAG ATTTGAATAT AAATGTATTT TTTAAACACT TGTTATGAGT 1680  
 TAACCTGGAT TACATTTTGA AATCAGTTCA TTCCATGATG CATATTACTG GATTAGATTA 1740  
 AGAAAGACAG AAAAGATTAA GGGACGGGCA CATTTTTCAA CGATTAAGAA TCATCATTAC 1800  
 15 ATAACCTGGT GAACTGAAA AAGTATATCA TATGGGTACA CAAGGCTATT TGCCAGCATA 1860  
 TATTAATATT TTAGAAAATA TTCTTTTGT AATACTGAAT ATAAACATAG AGCTAGAGTC 1920  
 ATATTATCAT ACTTATCATA ATGTTCAATT TGATACAGTA GAATTGCAAG TCCCTAAGTC 1980  
 CCTATTCAT GTGCTTAGTA GTGACTCCAT TTAATAAAAA GTGTTT TAG TTTTAAACAA 2040  
 CTAACCG

Seq ID NO: 101 Protein sequence:  
 Protein Accession #: NP\_000664

1 11 21 31 41 51  
 | | | | |  
 25 MG TAGKVIK KAAVLWEQK PFSIEEIEV PPKTKEVRIK ILATGICRTD DHVIKGTMV S 60  
 KFPVIVGHEA TGIVESIGEG VTTVKPGDKV IPLFLPQCRE CNACRNP DGN LCIRSDITGR 120  
 GVLADGTTTF TCKGKPVHMF MNTSTFT EYT VVDESSVAKI DDAAPEKVC LIGCGFSTGY 180  
 GAAVKTGKVK PGSTCVVFG L GVGLSVIMG CKSAGASRII GIDLNKDKFE KAMAVGATEC 240  
 30 ISPKDSTKPI SEVLSEMTGN NVGYTFE VIG HLETMIDALA SCHMNYGTSV VVGVP PPSAKM 300  
 LTYDPM L LFT GRTWKGC VFG GLKSRDDVPK LVTEFLAKKF DLDQLITHVL PFKKISEGFE 360  
 LLNSGQSIRT VLTF

Seq ID NO: 102 DNA sequence  
 Nucleic Acid Accession #: NM\_006783.1  
 Coding sequence: 1..786

1 11 21 31 41 51  
 | | | | |  
 40 ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAACACTC CACCAGCATC 60  
 GGAAGGTGT GGATCACAGT CATCTTTATT TTCCGAGTCA TGATCCTAGT GGTGGCTGCC 120  
 CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACCGGGATGC 180  
 AAAAATGTGT GCTATGACCA CTTTTTCCCG GTGTCCCACA TCCGGCTGTG GGCCCTCCAG 240  
 CTGATCTTCG TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCCTA CTACAGGCAC 300  
 45 GAAACCACTC GCAAGTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360  
 ATTA AAAAGC ACAAGGTT CG GATAGAGGGG TCGCTGTGGT GGACGTACAC CAGCAGCATC 420  
 TTTTCCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACAATGGG 480  
 TACCACCTGC CTTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCCAACCT GTTGTACTGC 540  
 TTTATTTCTA GGCCAACAGA GAAGACCGTG TTTACCATTT TTATGATTTT TCGTCTGTG 600  
 50 ATTTGCATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTT TAGG 660  
 AGATCAAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720  
 CAGAATGAAA TGAATGAGCT GATTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780  
 AGCTAA

Seq ID NO: 103 Protein sequence:  
 Protein Accession #: NP\_006774.1

1 11 21 31 41 51  
 | | | | |  
 60 MDWGLTLHFI GGVNKHSTSI GKWITVIFI FRVMILVVAA QEVWGDEQED FVCNTLQPGC 60  
 KNVCYDHFFP VSHIRLWALQ LIFVSTPALL VAMHVAYYRH ETTRKFRRGE KRNDFKDIED 120  
 IKKHKVRIEG SLWWTYTSSI FFRIIFEAAF MYVFYFLYNG YHLPWVLKCG IDPCPNLVDC 180  
 FISRPTEKTV FTIFMISASV ICMLLNVAEL CYLLLLKVCFR RSKRAQTQKN HPNHALKESK 240  
 QNEMNELISD SGQNAITGFP S

Seq ID NO: 104 DNA sequence  
 Nucleic Acid Accession #: NM\_020411  
 Coding sequence: 86-526

1 11 21 31 41 51  
 | | | | |  
 70 GGACCTGGGA AGGAGCATAG GACAGGGCAA GCGGGGATAA GGAGGGGCAC CACAGCCCTT 60  
 AAGGCACGAG GGAACCTCAC TGCGCATGCT CCTTTGGTGC CCACCTCAGT GCGCATGTTC 120  
 ACTGGGCGTC TTCCCATCGG CCCCTTCGCC AGTGTGGGGA ACGCGGCGGA GCTGTGAGCC 180  
 75 GGCGACTCGG GTCCCTGAGG TCTGGATTCT TTCTCCGCTA CTGAGACACG GCGGACACAC 240  
 ACAACACAG AACCACACAG CCAGTCCCAG GAGCCAGTA ATGGAGAGCC CCAAAAAGAA 300  
 GAACCAGCAG CTGAAAGTCG GGATCCTACA CCTGGGCAGC AGACAGAAGA AGATCAGGAT 360  
 ACAGCTGAGA TCCCAGTGCG CGACATGGAA GGTGATCTGC AAGAGCTGCA TCAGTCAAAC 420  
 ACCGGGGATA AATCTGGATT TGGGTTCCGG CGTCAAGGTG AAGATAATAC CTAAAGAGGA 480  
 80 ACCTGTAAA ATGCCAGAAG CAGGTGAAGA GCAACCACAA GTTTAAATGA AGACAAGCTG 540  
 AAACAACGCA AGCTGGTTTT ATATTAGATA TTTGACTTAA ACTATCTCAA TAAAGTTTGT 600  
 CAGCTTTCAC CAAAAA AAAA

Seq ID NO: 105 Protein sequence:  
 Protein Accession #: NP\_065144.1

1 11 21 31 41 51



MLLWCPPQCA CSLGVFSPAP SPVWGTRRSC EPATRVPEVW ILSPLLRHGG HTQTQNHAS 60  
PRSPVMESPK KKNQQLKVGI LHLGSRQKKI RIQLRSQCAT WKVICKSCIS QTPGINLDLG 120  
SGVKVKIIPK EEHCKMPEAG EEQPQV

Seq ID NO: 106 DNA sequence  
Nucleic Acid Accession #: J04129  
Coding sequence: 99-587

1 11 21 31 41 51  
| | | | | |  
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCTCCTGC 60  
TCACCCTGGG CGTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120  
AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CCTGGCACTC CATGGCCATG GCGACCAACA 180  
15 ACATCTCCCT CATGGCGACA CTGAAGGCC CTCTGAGGGT CCACATCACC TCACTGTTGC 240  
CCACCCCCGA GGACAACCTG GAGATCGTTC TGCACAGATG GGAGAACAAC AGCTGTGTTG 300  
AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG 360  
TGGCGAACGA GGCCACGCTG CTCGATACTG ACTACGACAA TTTCTGTGTT CTCTGCCTAC 420  
20 AGGACACCAC CACCCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCCCTGGTGG 480  
AGGACGATGA GATCATGCAG GGATTCATCA GGGCTTTCAG GCCCCTGCCC AGGCACCTAT 540  
GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCCGTGCCG TTTCTAGCTC ACCTCCGCCT 600  
CCAGGAAGAC CAGACTCCCA CCCTTCCACA CCTCCAGAGC AGTGGGACTT CCTCCTGCCC 660  
TTTCAAAGAA TAACCACAGC TCAGAAGACG ATGACGTGGT CATCTGTGTC GCCATCCCCCT 720  
25 TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCCTGGGGGC AGAGTCTCTG 780  
GCAGAGGTTA TTAATAAACC CTTGGAGCAT G

Seq ID NO: 107 Protein sequence:  
Protein Accession #: AAA60147

1 11 21 31 41 51  
| | | | | |  
MDIPQTKQDL ELPKLAGTWH SMAMATNNIS LMATLKAPLR VHITSLLPPT EDNLEIVLHR 60  
WENNSCVBKK VLGEKTGNPK KFKINYTVAN EATLLDSTDYD NFLFLCLQDT TTPIQSMCMQ 120  
35 YLARVLVEDD EIMQGFIRAF RPLPRHLWYL LDLQMEEPFC RF

Seq ID NO: 108 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 48-794

1 11 21 31 41 51  
| | | | | |  
TCCCAGGCAG CAGTTAGCCC GCCGCCCGCC TGTGTGTCCC CAGAGCCATG GAGAGAGCCA 60  
GTCTGATCCA GAAGGCCAAG CTGGCAGAGC AGGCCGAACG CTATGAGGAC ATGGCAGCCT 120  
45 TCATGAAAGG CGCCGTGGAG AAGGGCGAGG AGCTCTCCTG CGAAGAGCGA AACCTGCTCT 180  
CAGTAGCCTA TAAGAACGTG GTGGGCGGCT AGAGGGCTGC CTGGAGGGTG CTGTCCAGTA 240  
TTGAGCAGAA AAGCAACGAG GAGGGCTCGG AGGAGAAGGG GCCCGAGGTG CGTGAGTACC 300  
GGGAGAAGGT GGAGACTGAG CTCCAGGGCG TGTGCGACAC CGTGCTGGGC CTGCTGGACA 360  
50 GCCACCTCAT CAAGGAGGCC GGGGACGCCG AGAGCCGGGT CTTCTACCTG AAGATGAAGG 420  
GTGACTACTA CCGCTACCTG GCCGAGGTGG CCACCCGTGA CGACAAGAAG CGCATCATTG 480  
ACTCAGCCCC GTCCAGCCTAC CAGGAGGCCA TGGACATCAG CAAGAAGGAG ATGCCGCCCA 540  
CCAACCCCAT CCGCCTGGGC CTGGCCCTGA ACTTTCCGT CTTCCACTAC GAGATCGCCA 600  
ACAGCCCCGA GGAGGCCATC TCTCTGGCCA AGACCACTTT CGACGAGGCC ATGGCTGATC 660  
55 TGCACACCCT CAGCGAGGAC TCCTACAAAG ACAGCACCCT CATCATGCAG CTGCTGCGAG 720  
ACAACCTGAC ACTGTGGACG GCCGACAACG CCGGGGAAGA GGGGGGCGAG GCTCCCCAGG 780  
AGCCCCAGAG CTGAGTGTGG CCGGCCACCG CCCCGCCCTG CCCCCTCCAG TCCCCACCC 840  
TGCCGAGAGG ACTAGTATGG GGTGGGAGGC CCCACCCCTT TCCCCTAGGC GCTGTTCTTG 900  
CTCCAAAGGG CTCCGTGGAG AGGGACTGGC AGAGCTGAGG CCACCTGGGG CTGGGGATCC 960  
60 CACTCTTCTT GCAGCTGTTG AGCGCACCTA ACCACTGGTC ATGCCCCCAC CCCTGCTCTC 1020  
CGCACCCGCT TCCTCCCGAC CCCAGGACCA GGCTACTTCT CCCCTCCTCT TGCCTCCCTC 1080  
CTGCCCTGTC TGCCTCTGAT CGTAGGAATT GAGGAGTGTC CCGCCTTGTC GCTGAGAACT 1140  
GGACAGTGGC AGGGGCTGGA GATGGGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG 1200  
CGCGCGCGCC AGTGCAAGAC CGAGATTGAG GGAAAGCATG TCTGCTGGGT GTGACCATGT 1260  
65 TTCCTCTCAA TAAAGTTCCC CTGTGACACT C

Seq ID NO: 109 Protein sequence:  
Protein Accession #: NP\_006133.1

1 11 21 31 41 51  
| | | | | |  
MERASLIQKA KLAEQAERYE DMAAFMKGAV EKGEELSCEE RNLLSVAYKN VVGQRAAWR 60  
VLSSIEQKSN BEGSEKGPV VREYREKQET ELQGVCDTVL GLLDShLIKE AGDAESRVFY 120  
LKMKGDIYRY LAEVATGDDK KRIIDSARSA YQEAMDISKK EMPPTNPIRL GLALNFSVFH 180  
75 YEIANSPAEA ISLAKTTFDE AMADLHTLSE DSYKDSTLIM QLLRDNLTLW TADNAGEEGG 240  
EAPQEPQS

Seq ID NO: 110 DNA sequence  
Nucleic Acid Accession #: NM\_000695  
Coding sequence: 407-1564

1 11 21 31 41 51  
| | | | | |  
CACGAGTTGG TTTGGGAGCT GCCAGTCTCC TGGGAGGATC GCAGTCAGCA GAGCAGGGCT 60  
GAGGCCTGGG GGTAGGAGCA GAGCCTGCGC ATCTGGAGGC AGCATGTCCA AGAAAGGGAG 120  
85 TGGAGGTGCA GCGAAGGACC CAGGGGCGAG GCCCACGCTG GGGATGGACC CCTTCGAGGA 180  
CACACTGCGG CGGCTGCGTG AGGCCTTCAA CTGAGGGCGC ACGCGGCCGG CCGAGTTCCTG 240  
GGCTGCGCAG CTCCAGGGCC TGGGCCACTT CCTTCAAGAA AACAAGCAGC TTCTGCGCGA 300



	CGTGCTGGCC	CAGGACCTGC	ATAAGCCAGC	TTTCGAGGCA	GACATATCTG	AGCTCATCCT	360
	TTGCCAGAAC	GAGGTTGACT	ACGCTCTCAA	GAACCTTCAG	GCCTGGATGA	AGGATGAACC	420
	ACGGTCCACG	AACCTGTTCA	TGAAGCTGGA	CTCGGTCTTC	ATCTGGAAGG	AACCCCTTGG	480
5	CCTGGTCCTC	ATCATCGCAC	CCTGGAAC TA	CCCATTGAAC	CTGACCCTGG	TGCTCCTGGT	540
	GGGCACCC TC	CCCGCAGGGA	ATTGCGTG GT	GCTGAAGCCG	TCAGAAATCA	GCCAGGGCAC	600
	AGAGAAGGTC	CTGGCTGAGG	TGCTGCCCCA	GTACCTGGAC	CAGAGCTGCT	TTGCCGTGGT	660
	GCTGGGCGGA	CCCCAGGAGA	CAGGGCAGCT	GCTAGAGCAC	AAGTTGGACT	ACATCTTCTT	720
	CACAGGGAGC	CCTCGTGTGG	GCAAGATTGT	CATGACTGCT	GCCACCAAGC	ACCTGACGCC	780
10	TGTACCCCTG	GAGCTGGGGG	GCAAGAACCC	CTGCTACGTG	GACGACAACT	GCGACCCCCA	840
	GACCGTGGCC	AACCGCGTGG	CCTGGTTCTG	CTACTTCAAT	GCCGGCCAGA	CCTGCGTGGC	900
	CCCTGACTAC	GTCCTGTGCA	GCCCCGAGAT	GCAGGAGAGG	CTGCTGCCCG	CCCTGCAGAG	960
	CACCATCACC	CGTTTCTATG	GCGACGACCC	CCAGAGCTCC	CCAAACCTGG	GCCGCATCAT	1020
	CAACCAGAAA	CAGTTCCAGC	GGCTGCGGGC	ATTGCTGGGC	TGCGGCCGCG	TGGCCATTGG	1080
	GGGCCAGAGC	AACGAGAGCG	ATCGCTACAT	CGCCCCCAGC	GTGCTGGTGG	ACGTGCAGGA	1140
15	GACGGAGCCT	GTGATGCAGG	AGGAGATCTT	CGGGCCCATC	CTGCCCATCG	TGAACGTGCA	1200
	GAGCGTGGAC	GAGGCCATCA	AGTTTCATCA	CCGGCAGGAG	AAGCCCCCTG	CCCTGTACGC	1260
	CTTCTCCAAC	AGCAGACAGG	TTGTGAACCA	GATGCTGGAG	CGGACCAGCA	GCGGCAGCTT	1320
	TGGAGGCAAT	GAGGGCTTCA	CCTACATATC	TCTGCTGTCC	GTGCCATTCC	GGGGAGTCGG	1380
	CCACAGTGGG	ATGGGCCGGT	ACCACGGCAA	GTTACCTTTC	GACACCTTCT	CCCACCACCG	1440
20	CACCTGCCTG	CTCGCCCCCT	CCGGCCTGGA	GAAATTAAAG	GAGATCCGCT	ACCCACCCTA	1500
	TACCGACTGG	AACCAGCAGC	TGTTACGCTG	GGGCATGGGC	TCCCAGAGCT	GCACCCTCCT	1560
	GTGAGCGTCC	CACCCGCTTC	CAACGGGTCA	CACAGAGAAA	CCTGAGTCTA	GCCATGAGGG	1620
	GCTTATGCTC	CCAATCACA	TTGTTCTCTC	AGACCGCAGG	CTCCCCCAGC	CTCAGGTTGC	1680
25	TGGAGCTGTC	ACATGACTGC	ATCCTGCCTG	CCAGGGCTGC	AAAGCAAGGT	CTTGCTTCTA	1740
	TCTGGGGGAC	GCTGCTCGAG	AGAGGCCGAG	AGGCCGCAGA	ACATGCCAGG	TGTCCTCACT	1800
	CACCCACCCC	TCCCCAATTC	CAGCCCTTTG	CCCTCTCGGT	CAGGGTTGGC	CAGGCCCAGT	1860
	CACAGGGGCA	GTGTCACCC	GGAAAATACA	GTGCCCTGCC	TTCTTAGGGG	CATCAGCCCT	1920
	GAACGGTTGA	GAGCGTGGAG	CCCTCCAGGC	CTTTGCTCTC	CCCTCTAGGC	ACACGCGCAC	1980
30	TTCCACCTCT	GCCCCATCCC	AAC TGACCA	GCACTGCCTC	CCCCAGGGAT	CCTCTCACAT	2040
	CCCACACTGG	TCTCTGCACC	ACCCCTCTGG	TTACACCCGC	ACCCTGCACT	CACCCACAGC	2100
	AGTCCATCC	ACTGGGAAAA	CTGGGGTTTG	CATCACTCCA	CTGCACAGTG	TTAGTGGGAC	2160
	CTGGGGGCAA	GTCCCTTGAC	TTCTCTGAGC	CTCAGTTTCC	TTATGTGAAA	GTGCTGGAA	2220
	CCAAAATGGA	GTCAC TTATG	CCAACTCTA	ATAAAATGGA	GTGCGGGGGG	CACATAGAAG	2280
35	CCCTCACACA	CACATGCCCG	TAACAGGATT	TATCAACCAAG	ACACGCCTGC	ATGTAAGACC	2340
	AGACACAGGG	CGTATGAAA	AGCACGTCTT	CAAAGACTGT	AGTATTCCAG	ATGAGCTGCA	2400
	GATGCTTACC	TACCACGGCC	GTCTCCACCA	GAAAACCATC	GCCAACTCCT	GCGATCAGCT	2460
	TGTGACTTAC	AAACCTTGTT	TAAAAGCTGC	TTACATGGAC	TTCTGTCTCT	TAAAACGTTT	2520
40	CCCTTGCTG	TGGCCCTCTG	TGTATGCCTG	GGATCCTTCC	AAGCACTCAT	AGCCAGATA	2580
	GGAATCCTCT	GCTCCTCCCA	AATAAATTCA	TCTGTTC			

Seq ID NO: 111 Protein sequence:  
Protein Accession #: NP\_000686

45	1	11	21	31	41	51	
	MKDEPRSTNL	FMKLDSVFIW	KEPFGVLVII	APWNYPLNLT	LVLLVGTLP	GNCVVLKPSE	60
	ISQGTEKVL	EVLPQYLDQS	CFAVVLGGPQ	ETGQLLEHKL	DYIFFTGSPR	VGKIVMTAAT	120
50	KHLTPVTLE	GGKNPCYVDD	NCDPQTVANR	VAWFYFNAG	QTCVAPDYVL	CSPQMQRLL	180
	PALQSTITRF	YGDDPQSSPN	LGRIINQKQF	QRLRALLGCG	RVAIGGQSNE	SDRYIAPTIVL	240
	VDVQETEPVM	QEEIFGPILP	IVNVQSVDEA	IKFINRQKEP	LALYAFNSNR	QVVNQMLERT	300
	SSGSFGGNEG	FTYISLLSVP	FGGVGHSGMG	RYHGKFTFDT	FSHRTCLLA	PSGLEKLKEI	360
	RYPPYTDWNQ	QLLRWGMGSQ	SCTLL				

55 Seq ID NO: 112 DNA sequence  
Nucleic Acid Accession #: NM\_004456  
Coding sequence: 58-2298

60	1	11	21	31	41	51	
	GAATTCCGGG	CGACGCGCGG	GAACAACGCG	AGTCGGCGCG	CGGGACGAAG	AATAATCATG	60
	GGCCAGACTG	GGAAGAAATC	TGAGAAGGGA	CCAGTTTGTT	GGCGGAAGCG	TGTAAAATCA	120
	GAGTACATGC	GACTGAGACA	GCTCAAGAGG	TTCAGACGAG	CTGATGAAGT	AAAGAGTATG	180
65	TTTAGTTCCA	ATCGTCAGAA	AATTTTGGA	AGAACGGAAA	TCTTAAACCA	AGAATGGAAA	240
	CAGCGAAGGA	TACAGCCTGT	GCACATCTGT	ACTTCTGTGA	GCTCATTGCG	CGGGACTAGG	300
	GAGTGTTCCG	TGACCAAGTG	CTTGGAATTT	CCAACACAAG	TCATCCCAT	AAAGACTCTG	360
	AATGCAGTTG	CTTCAGTACC	CATAATGTAT	TCTTGGTCTC	CCCTACAGCA	GAATTTTATG	420
	GTGGAAGATG	AAACTGTTTT	ACATAACATT	CCTTATATGG	GAGATGAAGT	TTTAGATCAG	480
70	GATGTTACTT	TCATTGAAGA	ACTAATAAAA	AATTATGATG	GGAAAGTACA	CGGGGATAGA	540
	GAATGTGGGT	TTATAAATGA	TGAAATTTTT	GTGGAGTTGG	TGAATGCCCT	TGGTCAATAT	600
	AATGATGATG	ACGATGATGA	TGATGGAGAC	GATCCTGAAG	AAAGAGAAGA	AAAGCAGAAA	660
	GATCTGGAGG	ATCACCGAGA	TGATAAAGAA	AGCCGCCAC	CTCGGAAATT	TCCTTCTGAT	720
	AAAATTTTGG	AGGCCATTTT	CTCAATGTTT	CCAGATAAGG	GCACAGCAGA	AGAACTAAAG	780
75	GAAAAATATA	AAGAACTCAC	CGAACAGCAG	CTCCCAGGCG	CACTTCCTCC	TGAATGTACC	840
	CCCAACATAG	ATGGACCAAA	TGCTAAATCT	GTTTCAGAGAG	AGCAAAGCTT	ACACTCCTTT	900
	CATACGCTTT	TCTGTAGGCG	ATGTTTTTAA	TATGACTGCT	TCCTACATCC	TTTTCATGCA	960
	ACACCAACA	CTTATAAGCG	GAAGAACACA	GAAACAGCTC	TAGACAACAA	ACCTTGTGGA	1020
	CCACAGTGTT	ACCAGCATTT	GGAGGGAGCA	AAGGAGTTTG	CTGCTGCTCT	CACCGCTGAG	1080
80	CGGATAAAGA	CCCCACCAAA	ACGTCCAGGA	GGCCGCAGAA	GAGGACGGCT	TCCAATAAAC	1140
	AGTAGCAGGC	CCAGCACCCC	CACCATTAA	TGCTGGAAT	CAAAGGATAC	AGACAGTGAT	1200
	AGGGAAGCAG	GGACTGAAAC	GGGGGGAGAG	AACAATGATA	AAGAAGAAGA	AGAGAAGAAA	1260
	GATGAAACTT	CGAGCTCCTC	TGAAGCAAAT	TCTCGGTGTC	AAACACCAAT	AAAGATGAAG	1320
	CCAAATATTG	AACCTCCTGA	GAATGTGGAG	TGGAGTGGTG	CTGAAGCCTC	AATGTTTAGA	1380
85	GTCTCATTTG	GCACTTACTA	TGACAATTTT	TGTGCCATTG	CTAGGTTAAT	TGGGACCAAA	1440
	ACATGTAGAC	AGGTGTATGA	GTTTATAGTC	AAAGAATCTA	GCATCATAGC	TCCAGCTCCC	1500
	GCTGAGGATG	TGGATACTCC	TCCAAGGAAA	AAGAAGAGGA	AACACCGGTT	GTGGGCTGCA	1560
	CACTGCAGAA	AGATACAGCT	GAAAAAGGAC	GGCTCCTCTA	ACCATGTTTA	CAACTATCAA	1620



	CCCTGTGATC	ATCCACGGCA	GCCTTGTGAC	AGTTCGTGCC	CTTGTGTGAT	AGCACAAAAT	1680
	TTTTGTGAAA	AGTTTGTCA	ATGTAGTTCA	GAGTGTCAAA	ACCGCTTTCC	GGGATGCCGC	1740
	TGCAAAGCAC	AGTGCAACAC	CAAGCAGTGC	CCGTGCTACC	TGGCTGTCCG	AGAGTGTGAC	1800
5	CCTGACCTCT	GTCTTACTTG	TGGAGCCGCT	GACCATTGGG	ACAGTAAAAA	TGTGTCTCTG	1860
	AAGAACTGCA	GTATTCAGCG	GGGCTCCAAA	AAGCATCTAT	TGCTGGCACC	ATCTGACGTG	1920
	GCAGGCTGGG	GGATTTTTAT	CAAAGATCCT	GTGCAGAAAA	ATGAATTCAT	CTCAGAATAC	1980
	TGTGGAGAGA	TTATTTCTCA	AGATGAAGCT	GACAGAAGAG	GGAAAGTGTA	TGATAAATAC	2040
	ATGTGCAGCT	TTCTGTTCAG	CTTGAACAAT	GATTTTGTGG	TGGATGCAAC	CCGCAAGGGT	2100
10	AACAAAATTC	GTTTTGCAAA	TCATTCGGTA	AATCCAAACT	GCTATGCAAA	AGTTATGATG	2160
	GTAAACGGTG	ATCACAGGAT	AGGTATTTTT	GCCAAGAGAG	CCATCCAGAC	TGGCGAAGAG	2220
	CTGTTTGTG	ATTACAGATA	CAGCCAGGCT	GATGCCCTGA	AGTATGTCGG	CATCGAAAGA	2280
	GAAATGGAAA	TCCCTTGACA	TCTGCTACCT	CCTCCCCCTC	CTCTGAAACA	GCTGCCTTAG	2340
	CTTCAGGAAC	CTCGAGTACT	GTGGGCAATT	TAGAAAAAGA	ACATGCAGTT	TGAAATTCTG	2400
15	AATTTGCAAA	GTAAGTAAAG	AATAATTTAT	AGTAATGAGT	TTAAAAATCA	ACTTTTTATT	2460
	GCCTTCTCAC	CAGCTGCAAA	GTGTTTTGTA	CCAGTGAATT	TTTGCAATAA	TGCAGTATGG	2520
	TACATTTTTC	AACTTTGAAT	AAAGAATACT	TGAAC TTGAA	AAAAAAAAAA	AAAAAA	

20 Seq ID NO: 113 Protein sequence:  
Protein Accession #: NP\_004447

	1	11	21	31	41	51	
25	MGQTGKKSEK	GPVCWRKRVK	SEYMRLRQLK	RFRRADDEVKS	MFSSNRQKIL	ERTEILNQEW	60
	KQRRIQPVHI	LTSVSSLRGT	RECSVTSDDL	FPTQVIPLKT	LNAVASVPIM	YSWSPLQQNF	120
	MVEDETVLHN	IPYMGDEVLD	QDGTFFIEELI	KNYDGVHGD	RECGFINDEI	FVELVNALGQ	180
	YNDDDDDDDG	DDPEEREKQ	KDLEDHRDDK	ESRPPRKFPS	DKILEAISSM	FPDKGTAEEL	240
	KEKYKELTEQ	QLPGALPPEC	TPNIDGPNK	SVQREQSLHS	FHTLFCRRCF	KYDCFLHPFH	300
30	ATPNTYKRKN	TETALDNKPC	GPQCYQHLEG	AKEFAAALTA	ERIKTPPKRP	GGRRRGRLPN	360
	NSSRPSTPTI	NVLESKDTDS	DREAGTETGG	ENNDKEEEK	KDETSSSSEA	NSRCQTPIKM	420
	KPNIEPPENV	EWSGAEASMF	RVLIGTYIDN	FCAIARLIGT	KTCRQVYEF	VKESSIIAPA	480
	PAEDVDTPPR	KKKRKHRLWA	AHCRKIQLKK	DGSSNHVYNY	QPCDHPRQPC	DSSPCPCVIAQ	540
	NFCEKFCQCS	SECQNRFPGC	RCKAQCNKTK	CPCYLAVREC	DPDLCLTCGA	ADHWDSKNVS	600
35	CKNCSIQRGS	KKHLLAPSD	VAGWIFIKD	PVQKNEFISE	YCGEIIISQDE	ADRRGKVYDK	660
	YMCSFLFNLN	NDFVVDATRK	GNKIRFANHS	VNPNCYAKVM	MVNGDHRIGI	FAKRAIQTGE	720
	ELFVDYRYSQ	ADALKYVGIE	REMEIP				

40 Seq ID NO: 114 DNA sequence  
Nucleic Acid Accession #: NM\_001827  
Coding sequence: 96-335

	1	11	21	31	41	51	
45	AGTCTCCGGC	GAGTTGTTCG	CTGGGCTGGA	CGTGGTTTTG	TCTGCTGCGC	CCGCTCTTCG	60
	CGCTCTCGTT	TCATTTTCTG	CAGCGCGCCA	CGAGGATGGC	CCACAAGCAG	ATCTACTACT	120
	CGGACAAGTA	CTTCGACGAA	CACCTACGAGT	ACCGGCATGT	TATGTTACCC	AGAGAACCTT	180
	CCAAACAAGT	ACCTAAAACT	CATCTGATGT	CTGAAGAGGA	GTGGAGGAGA	CTTGGTGTCC	240
	AACAGAGTCT	AGGCTGGGTT	CATTACATGA	TTATGAGGCC	AGAACCACAT	ATTCTTCTCT	300
50	TTAGACGACC	TCTTCCAAAA	GATCAACAAA	AATGAAGTTT	ATCTGGGGAT	CGTCAAATCT	360
	TTTTCAAATT	TAATGTATAT	GTGTATATAA	GGTAGTATTC	AGTGAATACT	TGAGAAATGT	420
	ACAAATCTTT	CATCCATACC	TGTGCATGAG	CTGTATTCTT	CACAGCAACA	GAGCTCAGTT	480
	AAATGCAACT	GCAAGTAGGT	TACTGTAAAG	TGTTTAAGAT	AAAAGTTCTT	CCAGTCAGTT	540
55	TTTCTCTTAA	GTGCCTGTTT	GAGTTTATCG	AAACAGTTTA	CTTTTGTTC	ATAAAGTTTG	600
	TATGTTGCAT	TTAAAAAAA	AAAAAA				

Seq ID NO: 115 Protein sequence:  
Protein Accession #: NP\_001818

	1	11	21	31	41	51	
60	MAHKQIYYSD	KYFDEHYEYR	HVMLPRELSK	QVPKTHLMSE	BEWRRRLGVQQ	SLGWVHYMIH	60
	EPEPHILLFR	RPLPKDQK					

65 Seq ID NO: 116 DNA sequence  
Nucleic Acid Accession #: CAT cluster

	1	11	21	31	41	51	
70	TCAGACCTCA	TGAGTCACTT	GGACTCTTGA	GCCACCTCTG	GGGGTGGAGT	CTCTCTCCTG	60
	GCATCTGGAC	CCTTGGTGCT	ATCGACGAAG	CTTGGGTGGG	GCTCTTAGCT	GCTATGTGCA	120
	AGAGGTGTGT	TCCAGGGAAA	GCCCCATCT	CTCTGCAGAG	GTCAAGTGAA	AGCGACGGCC	180
	GCAGCCAACA	GAGTTCAAAA	TGCAGGCTTG	GAAAGTACAG	GGGGCTCTGT	GGAGGATGGG	240
	AAGGACTGAT	CCACATTCCC	ACCAGGAAGT	TTAGCAGAAC	CCCCGCGTGC	CAACTGGACC	300
75	CCTTGGGAAG	ACCTGGCTCA	GGCTGGACCA	CCTCTTGAGA	GGGAGGAGCT	CTGGATTGTA	360
	TCAAGAATTC	TTTGCTGAGC	ATGGTGCTTC	ATGCCTATAA	TACCAACACT	TTGGGAGGCC	420
	AGTGTGGGAG	GATCTCTTGA	GCCCAGGAGT	TCAAGACTAG	CCTGGGCAAC	ACAGAGAGAA	480
	CCCATCTCTA	AAATAATAAT	AATAATAAAA	TAAAAAATTA	GCAGGGCATG	GTGGCATGTG	540
	CCTGTAGTTC	CAGCTACCCA	GGAGGCTGAG	GCAAGAGGAT	GGCTGGAGCC	TGGGATGTTG	600
80	AGGCTGCAAT	GAACTGTGAT	TACCCCACTG	CACTCCAGCC	TGGGCAAAAG	AGCGAGAGAA	660
	CCTGTCTCAA	ATAATAATAA	TAATAATAAT	CTTATTTTGG	AGAATAAAGA	GACCTCTGGA	720
	TTTGAGGTGC	CATTTGGGTA	GAAAGAAAAG	ACGTTTACAC	CGAGAAATAG	TCTGTGTTGC	780
	CCTGAAGGAG	CAGAGGGATG	CATCGCTGGA	GGTGACCTAC	AGTTGAAGAA	GACTCATTAT	840
	GACGACCTT	GTCTTCTTTC	CTTGTGGAAA	GTGTTTCTCT	TGCTGCTACT	GCTCATGAGA	900
85	CTCTTCCCCC	TCCCTGTCCC	AGGGGAACAA	AGGGCTTTCT	ACCACACCCT	TTCTTGCCCC	960
	CCGCTCCCA	TGTCTGCTGT	GCCTTTGTAC	TCAGCAATTC	TTGTTTGCTC	CATTATCTTC	1020
	CAGCCGGATA	CAGAGTGAAT	AGTTAACCAC	ACTTAGGTCA	AATAGGATCT	AAATTTTTGT	1080
	TCCTGCTCCG	TGTAAAGAGG	CCAGTGTGTT	TGTGTTGCAA	GCAGCCTTGG	AATAGTAAC	1140



CTTCTCATTT GTTTGGGATC TGGCCACCAA GTTCCAGAA TATACACGGA TCAGTGCAGA 1200  
 AGTTTCATCAG GCTCTCGGAC CTTAGGGCTG TTGGAGAAGG CTTACAGCAGC AGAACTGATG 1260  
 GTGAAGGCTC GTGTTCTCCA TCCTCAACTT TCTTTGCTTC GATCATAAC AAGAATACAT 1320  
 TTGGAAGGGC AAAAAATGAA CACTGTCTGT CATTGCAGCC GTGTTTGTG ACACAGATGC 1380  
 ACAGTCTGCT GTGAAGACCT TCTCTCAAGT GGCATTTGGG AGTCCATGCC AGATCATGGT 1440  
 GCTTCATGAG AGACTGACAG CTATCAGGGG TTGTGGCACT TAGTGAGGAC TCTCCTCCCC 1500  
 CAGTGTGTGC TGATGACACA TACACACCTG ACAATAGCTT GAGTCTTCTC TGTTCTTTT 1560  
 ACTCTGTAGC CAACATACAC ATGATTTAAA ACCCTTTCTA AATATCTATC ATGGTTCATC 1620  
 CTTGTCCAAA TGCAGAGTCA GAGCTATTTG TACTTCATTA TTATTTCCAA GGCGAATAGT 1680  
 TGGCTTTCTT TTTGCAAAAA TAATTAAAGT TTTTGTATGT TGCAAAAAAA AAAAAAATAA 1740  
 AAACAAAAAA

Seq ID NO: 117 DNA sequence  
 Nucleic Acid Accession #: BC012178.1  
 Coding sequence: 204-2285

1 11 21 31 41 51  
 | | | | |  
 CTTCTCTCCC GCGGCGCTGG GCGCCGCGCT CCGCTGCTGT TGCTCCATTC GGCGCTTTTC 60  
 TGGCGGCTGG CTCTCTCCG CTGCGGCTG CTCCTCGACC AGGCTCCTT CTCAACCTCA 120  
 GCGCGCGGCG CCGACCTTC CCGCACCTC CCGCCCGCTC TCGTACTGTC GCCGTCACCG 180  
 CCGCGGCTCC GGCCCTGGCC CCGATGGCTC TGTGCAACGG AGACTCCAAG CTGGAGAATG 240  
 CTGGAGGAGA CCTAAGGAT GGCCACCACC ACTATGAAGG AGCTGTTGTC ATTCTGGATG 300  
 CTGGTGCTCA GTACGGGAAA GTCATAGACC GAAGAGTGAG GGAAGTGTTC GTGCAGTCTG 360  
 AAATTTTCCC CTTGGAAACA CCAGCATTTG CTATAAAGGA ACAAGGATTC CGTGCTATTA 420  
 TCATCTCTGG AGGACCTAAT TCTGTGTATG CTGAAGATGC TCCCTGGTTT GATCCAGCAA 480  
 TATTCATCTAT TGGCAAGCCT GTTCTTGGAA TTTGCTATGG TATGCAGATG ATGAATAAGG 540  
 TATTTGGAGG TACTGTGCAC AAAAAAAGTG TCAGAGAAGA TGGAGTTTTC AACATTAGTG 600  
 TGGATAATAC ATGTTTCTTA TTCAGGGGCC TTCAGAAGGA AGAAGTTGTT TTGCTTACAC 660  
 ATGGAGATAG TGTAGACAAA GTAGCTGATG GATTCAAGGT TGTGGCACGT TCTGGAAACA 720  
 TAGTAGCAGG CATAGCAAA TGAATCTAAA AGTTATATGG AGCACAGTTC CACCCTGAAG 780  
 TTGGCCTTAC AGAAAAATGA AAAGTAATAG TGAAGAATT CTTTATGAT ATAGCTGGAT 840  
 GCAGTGGAAC CTTACCGTG CAGAAACAG AACTTGAGTG TATTCGAGAG ATCAAAGAGA 900  
 GAGTAGGCAC GTCAAAAGTT TTGGTTTAC TCAGTGGTGG AGTAGACTCA ACAGTTTGTA 960  
 CAGCTTTGCT AAATCGTGCT TTGAACCAAG AACAAGTCAT TGCTGTGCAC ATTGATAATG 1020  
 GCTTTATGAG AAAACGAGAA AGCCAGTCTG TTGAAGAGGC CCTCAAAAAG CTTGGAATTC 1080  
 AGGTCAAAGT GATAAATGCT GCTCATTCTT TCTACAATGG AACAACAACC CTACCAATAT 1140  
 CAGATGAAGA TAGAACCCCA CGGAAAAGAA TTAGCAAAAC GTTAAATATG ACCACAAGTC 1200  
 CTGAAGAGAA AAGAAAAATC ATTGGGGATA CTTTTGTTAA GATTGCCAAT GAAGTAATTG 1260  
 GAGAAATGAA CTTGAAACCA GAGGAGGTTT TCCTTGCCCA AGGTACTTTA CGGCCTGATC 1320  
 TAATTGAAAG TGCATCCCTT GTTGCAAGTG GCAAAGCTGA ACTCATCAA ACCCATCACA 1380  
 ATGACACAGA GCTCATCAGA AAGTTGAGAG AGGAGGGAAA AGTAATAGAA CCTCTGAAAG 1440  
 ATTTTCATAA AGATGAAGTG AGAATTTTGG GCAGAGAACT TGGACTTCCA GAAGAGTTAG 1500  
 TTTCCAGGCA TCCATTTCCA GGTCTTGGCC TTGCAATCAG AGTAATATGT GCTGAAGAAC 1560  
 CTTATATTTG TAAGGACTTT CTTGAAACCA ACAATATTTT GAAAATAGTA GCTGATTTTT 1620  
 CTGCAAGTGT TAAAAAGCCA CATACCTAT TACAGAGAGT CAAAGCCTGC ACAACAGAAG 1680  
 AGGATCAGGA GAAGCTGATG CAAATTACCA GTCTGCATTC ACTGAATGCC TTCTTGCTGC 1740  
 CAATTAAAC TGTAGGTGTG CAGGGTGACT GTCGTTCCCTA CAGTTACGTG TGTGGAATCT 1800  
 CCAGTAAAGA TGAACCTGAC TGGGAATCAC TTATTTTCTT GGCTAGGCTT ATACCTCGCA 1860  
 TGTGTACAA CGTTAACAGA GTTGTTTATA TATTTGGCCC ACCAGTTAAA GAACCTCCTA 1920  
 CAGATGTTAC TCCCACTTTC TTGACAACAG GGTGCTCAG TACTTTACGC CAAGCTGATT 1980  
 TTGAGGCCCA TAACATCTC AGGAGTCTG GGTATGCTGG GAAAATCAGC CAGATGCCGG 2040  
 TGATTTTGAC ACCATTACAT TTTGATCGGG ACCCACTTCA AAAGCAGCCT TCATGCCAGA 2100  
 GATCTGTGGT TATTCGAACC TTTATTACTA GTGACTTCAT GACTGGTATA CCTGCAACAC 2160  
 CTGGCAATGA GATCCCTGTA GAGGTGGTAT TAAAGATGGT CACTGAGATT AAGAAGATTC 2220  
 CTGGTATTTT TCGAATTATG TATGACTTAA CATCAAAGCC CCCAGGAACT ACTGAGTGGG 2280  
 AGTAATAAAC TTCTTGTCTT ATTAATAA

Seq ID NO: 118 Protein sequence:  
 Protein Accession #: AAH12178.1

1 11 21 31 41 51  
 | | | | |  
 MALCNGDSKL ENAGGDLKDG HHHYEGAVVI LDAGAQYGVV IDRRVRELTV QSEIFPLETP 60  
 AFAIKEQGFR AIIISGGPNS VYAEADAPWF PAIFTIGKPV LGICYGMQMM NKVFGGTVHK 120  
 KSVREDGVFN ISVDNCTSLF RGLQKEEVVL LTHGDSVDKV ADGFKVVARV GNIVAGIANE 180  
 SKKLYGAQFH PEVGLTENGL VILKNFLYDI AGCSGTFTVQ NRELECIRES KERVGTSKVL 240  
 VLLSGGVDSV VCTALLNRAL NQEQVIAVHI DNGFMRKRES QSVVEALKKL GIQVKVINAA 300  
 HSFYNGTTTL PISDEDRTPR KRISKTLNMT TSPEEKRII GDTFVKIANE VIGEMNLKPE 360  
 EVFLAQGTLR PDLIESASLV ASGKAELIKT HHNDTELIRK LREEGVIEP LKDFHKDEV 420  
 ILGRELGLPE ELVSRHPFPG PGLAIRVICA EEPYICKDFP ETNNILKIVA DFSASVKKPH 480  
 TLLQVRVACT TEEDQEKLMQ ITSLHSLNAF LLPIKTVGVQ GDCRSYSYVC GISSKDEPDW 540  
 ESLIFLARLI PRMCHNVNRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADFEAHNILR 600  
 ESGYAGKISQ MPVILTPLHF DRDPLQKQPS QRSVVIRTG ITSDFMTGIP ATPGNEIPVE 660  
 VVLKMTVEIK KIPGISRIMY DLTSKPPGTT EWE

Seq ID NO: 119 DNA sequence  
 Nucleic Acid Accession #: NM\_006500.1  
 Coding sequence: 27..1967

1 11 21 31 41 51  
 | | | | |  
 ACTTGCCTCT CGCCCTCCGG CCAAGCATGG GGCTTCCAG GCTGGTCTGC GCCTTCTTGC 60  
 TCGCCGCTG CTGCTGCTGT CCTCGCGTGG CCGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120  
 CGCCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCTCTCTCC 180  
 AGTCCAAGG CAACCTCAGC CATGTGCGACT GGTTTTCTGT CCACAAGGAG AAGCGGACGC 240



	TCATCTTCCG	TGTGCGCCAG	GGCCAGGGCC	AGAGCGAACC	TGGGGAGTAC	GAGCAGCGGC	300
	TCAGCCTCCA	GGACAGAGGG	GCTACTCTGG	CCCTGACTCA	AGTCACCCCC	CAAGACGAGC	360
	GCATCTTCTT	GTGCCAGGGC	AAGCGCCCTC	GGTCCCAGGA	GTACCGCATC	CAGCTCCGCG	420
5	TCTACAAAGC	TCCGGAGGAG	CCAAACATCC	AGGTCAACCC	CCTGGGCATC	CCTGTGAACA	480
	GTAAGGAGCC	TGAGGAGGTC	GCTACCTGTG	TAGGGAGGAA	CGGGTACCCC	ATTCTCTAAG	540
	TCATCTGGTA	CAAGAATGGC	CGGCCTCTGT	AGGAGGAGAA	GAACCGGGTC	CACATTCACT	600
	CGTCCCAGAC	TGTGGAGTCG	AGTGGTTTGT	ACACCTTGCA	GAGTATTCTG	AAGGCACAGC	660
	TGGTTAAAGA	AGACAAAGAT	GCCCAGTTTT	ACTGTGAGCT	CAACTACCGG	CTGCCCAGTG	720
10	GGAACCACAT	GAAGGAGTCC	AGGGAAGTCA	CCGTCCCTGT	TTTCTACCCG	ACAGAAAAAG	780
	TGTGGCTGGA	AGTGGAGCCC	GTGGGAATGC	TGAAGGAAGG	GGACCGCGTG	GAAATCAGGT	840
	GTTTGGCTGA	TGGCAACCC	CCACCACACT	TCAGCATCAG	CAAGCAGAAC	CCCAGCACCA	900
	GGGAGGCAGA	GGAAGAGACA	ACCAACGACA	ACGGGGTCCT	GGTGCTGGAG	CCTGCCCCGA	960
	AGGAACACAG	TGGGCGCTAT	GAATGTCAAG	CCTGGAACTT	GGACACCATG	ATATCGCTGC	1020
15	TGAGTGAACC	ACAGGAACCTA	CTGGTGAAC	ATGTGTCTGA	CGTCCGAGTG	AGTCCCGCAG	1080
	CCCCTGAGAG	ACAGGAAGGC	AGCAGCCTCA	CCCTGACCTG	TGAGGCAGAG	AGTAGCCAGG	1140
	ACCTCGAGTT	CCAGTGGCTG	AGAGAAGAGA	CAGACCAGGT	GCTGGAAAGG	GGGCCTGTGC	1200
	TTCAGTTGCA	TGACCTGAAA	CGGGAGGCAG	GAGGCGGCTA	TCGCTGCGTG	GCGTCTGTGC	1260
	CCAGCATACC	CGGCCTGAAC	CGCACACAGC	TGGTCAAGCT	GGCCATTTTT	GGCCCCCTTT	1320
20	GGATGGCATT	CAAGGAGAGG	AAGGTGTGGG	TGAAAGAGAA	TATGGTGTG	AATCTGTCTT	1380
	GTGAAGCGTC	AGGGCAGCCC	AGGGCCACCA	TCTCCTGGAA	CGTCAACGGC	ACGGCAAGTG	1440
	AACAAGACCA	AGATCCACAG	CGAGTCCTGA	GCACCTGAA	TGTCCTCGTG	ACCCCGGAGC	1500
	TGTTGGAGAC	AGGTGTTGAA	TGCACGGCCT	CCAACGACCT	GGGCAAAAC	ACCAGCATCC	1560
	TCTTCCTGGA	GCTGGTCAAT	TTAACCACCC	TCACACCAGA	CTCCAACACA	ACCACTGGCC	1620
25	TCAGCACTTC	CACTGCCAGT	CCTCATACCA	GAGCCAACAG	CACCTCCACA	GAGAGAAAGC	1680
	TGCCGGAGCC	GGAGAGCCGG	GGCGTGGTCA	TCGTGGCTGT	GATTGTGTGC	ATCCTGGTCC	1740
	TGGCGGTGCT	GGGCGCTGTC	CTCTATTCC	TCTATAAGAA	GGGCAAGCTG	CCGTGCAGGC	1800
	GCTCAGGGAA	GCAGGAGATC	ACGCTGCCCC	CGTCTCGTAA	GACCGAACTT	GTAGTTGAAG	1860
	TTAAGTCAGA	TAAGCTCCCA	GAAGAGATGG	GCCTCCTGCA	GGGCAGCAGC	GGTGACAAGA	1920
30	GGGCTCCGGG	AGACCAGGGA	GAGAAATACA	TCGATCTGAG	GCATTAGCCC	CGAATCACTT	1980
	CAGCTCCCTT	CCCTGCCCTG	ACCATTCCTA	GCTCCCTGCT	CACTCTTCTC	TCAGCCAAAG	2040
	CCTCCAAAGG	GACTAGAGAG	AAGCCTCCTG	CTCCCTCCTC	CTGCACACCC	CCTTTCAGAG	2100
	GGCCACTGGG	TTAGGACCTG	AGGACCTCAC	TTGGCCCTGC	AAGCCGCTTT	TCAGGGACCA	2160
	GTCCACCACC	ATCTCCTCCA	CGTTGAGTGA	AGCTCATCCC	AAGCAAGGAG	CCCCAGTCTC	2220
35	CCGAGCGGGT	AGGAGAGTTT	CTTGAGAAC	GTGTTTTTTC	TTTACACACA	TTATGGCTGT	2280
	AAATACCTGG	CTCCTGCCAG	CAGCTGAGCT	GGGTAGCCTC	TCTGAGCTGG	TTTCTGCCC	2340
	CAAAGGCTGG	CTTCCACCAT	CCAGGTGCAC	CACTGAAGTG	AGGACACACC	GGAGCCAGGC	2400
	GCCTGCTCAT	GTTGAAGTGC	GCTGTTTACA	CCCGCTCCGG	AGAGCACCCC	AGCGGCATCC	2460
	AGAAGCAGCT	GCAGTGTGTC	TGCCACCACC	CTCCTGCTCG	CCTCTTCAAA	GTCTCCTGTG	2520
40	ACATTTTTTC	TTTGGTCAGA	AGCCAGGAAC	TGGTGTGATT	CCTTAAAAGA	TACGTGCCGG	2580
	GGCCAGGTGT	GGTGGCTCAC	GCCTGTAATC	CCAGCACTTT	GGGAGGCCGA	GGCGGGCGGA	2640
	TCACAAAGTC	AGGACGAGAC	CATCCTGGCT	AACACGGTGA	AACCTGTCT	CTACTAAAAA	2700
	TACAAAAAAA	AATTAGCTAG	GCGTAGTGGT	TGGCACCTAT	AGTCCAGCT	ACTCGGAAGG	2760
	CTGAAGCAGG	AGAATGGTAT	GAATCCAGGA	GGTGGAGCTT	GCAGTGAGCC	GAGACCGTGC	2820
45	CAGTGCATC	CAGCCTGGGC	AACACAGCGA	GACTCCGTCT	CGAGGAAAAA	AAAAGAAAAG	2880
	ACCGGTACCT	GCGGTGAGGA	AGCTGGGCGC	TGTTTTCGAG	TTCAAGTGAA	TTAGCCTCAA	2940
	TCCCCGTGTT	CACCTGCTCC	CATAGCCCTC	TTGATGGATC	ACGTAAAACT	GAAAGGCAGC	3000
	GGGGAGCAGA	CAAAGATGAG	GTCTACACTG	TCCTTCATGG	GGATTAAAGC	TATGGTTATA	3060
	TTAGCACCAA	ACTTCTACAA	ACCAAGCTCA	GGGCCCCAAC	CCTAGAAGGG	CCCAAATGAG	3120
50	AGAATGGTAC	TTAGGGATGG	AAAACGGGGC	CTGGCTAGAG	CTTCGGGTGT	GTGTGTCTGT	3180
	CTGTGTGTAT	GCATACATAT	GTGTGTATAT	ATGTTTTTGT	CAGGTGTGTA	AATTTGCAAA	3240
	TTGTTTTCTT	TATATATGTA	TGTATATATA	TATATGAAAA	TATATATATA	TATGAAAAAT	3300
	AAAGCTTAAT	TGTCCCAGAA	AATCATACAT	TGCTTTTTTA	TTCTACATGG	GTACCACAGG	3360
	AACCTGGGGG	CCTGTGAAAC	TACAACCAAA	AGGCACACAA	AACCGTTTCC	AGTTGGCAGC	3420
55	AGAGATCAGG	GGTTACCTCT	GCTTCTGAGC	AAATGGCTCA	AGCTCTACCA	GAGCAGACAG	3480
	CTACCCTACT	TTTCAGCAGC	AAAACGTCCC	GTATGACGCA	GCACGAAGGG	CCTGGCAGGC	3540
	TGTTAGCAGG	AGCTATGTCC	CTTCCTATCG	TTTCCGTCCA	CTT		

Seq ID NO: 120 Protein sequence:  
Protein Accession #: NP\_006491.1

	1	11	21	31	41	51	
65	MGLPRLVCAF	LLAACCCCP	VAGVPGEAEQ	PAPELVEVEV	GSTALLKCG	SQSQGNLSHV	60
	DWFSVHKEKR	TLIFRVRQGG	GQSEPGYEYQ	RLSLQDRGAT	LALTQVTPQD	ERIFLCQGKR	120
	PRSQEYRIQL	RVYKAPBEPN	IQVNPLGIPV	NSKEPEEVAT	CVGRNGYPIP	QVIWYKNGRP	180
	LKEEKNRVHI	QSSQTVESSE	LYTLQSLILK	QLVKEDKDAQ	FYCELNYRLP	SGNHMKESRE	240
70	VTVPVFYPT	KVWLEVEPVG	MLKEGDRVEI	RCLADGNPPP	HFSISKQNP	TREAEETT	300
	DNGVLVLEPA	RKEHSGRYEC	QAWNLDTMIS	LLSEPQELLV	NYVSDVRVSP	AAPERQEGSS	360
	LTLTCEAESS	QDLEFQWLRE	ETDQVLERGP	VLQLHDLKRE	AGGGYRCVAS	VPSIPGLNRT	420
	QLVKLAIFGP	PWMAFKERKV	WVKENMVLNL	SCEASGHPRP	TISWNVNGTA	SEQDQDPQRV	480
	LSTLNLVLT	ELLETGVECT	ASNDLGKNTS	ILFLELVNLT	TLTPDSNTTT	GLSTSTASPH	540
75	TRANSTSTER	KLPEPESRGV	VIVAVIVCIL	VLAVLGAVLY	FLYKKGKLPC	RRSGKQEITL	600
	PPSRKTELTV	EVKSDKLPEE	MGLLQGS	KRAPGDQGEK	YIDLRLH		

Seq ID NO: 121 DNA sequence  
Nucleic Acid Accession #: NM\_018306  
Coding sequence: 60-671

	1	11	21	31	41	51	
80	ATAGTCTACA	CAGAGCTCCC	CTTGCTGCCC	AGACAAGCTG	AAGGACCACA	GGAAAAGCCA	60
	TGGAGACTTC	AGCATCCTCC	TCCCAGCCTC	AGGACAACAG	TCAAGTCCAC	AGAGAAACAG	120
85	AAGATGTAGA	CTATGGAGAG	ACAGATTTC	ACAAGCAAGA	CGGGAAGGCT	GGACTCTTTT	180
	CCCAAGAACA	ATATGAGAGA	AACAAGTCTT	CTTCCTCCTC	CTTCTCTTCC	TCCTCATCCT	240
	CCTCATCTTC	TTCATCCTCC	TCCTCCTCAG	GTCCTGGGCA	TGGGGAGCCT	GACGTTTTGA	300



5  
10  
15  
20

```
AGGATGAGCT TCAACTCTAT GGAGATGCTC CTGGAGAGGT GGTACCCTCT GGGGAATCAG 360
GACTCCGAAG GAGAGGCTCT GACCCAGCAA GTGGAGAAGT GGAGGCCTCT CAGTTAAGAA 420
GACTGAATAT AAAGAAAGAT GATGAGTTTT TCCATTTCTG CCTCCTGTGC TTTGCCATCG 480
GGGCCTTGCT GGTGTGTTAT CACTATTACG CAGACTGGTT CATGTCCTTT GGGGTCGGCC 540
TGCTCACCTT CGCCTCCCTG GAAACCGTTG GCATCTACTT CGGACTAGTG TACCGTATCC 600
ACAGCGTCCT CCAAGGCTTC ATCCCCCTCT TCCAGAAGTT TAGGCTGACA GGGTTCAGGA 660
AGACTGACTG AGGCCACTTC CAGGTGGGCA GCAGAGGCAG GCCCCAGTGT GACCACCACT 720
GCGACCCCTG AGCCCAACAAG GGCAGAGCAG CATTCTGAGA GACGCACAGG AGACCAAGCC 780
AGACCAATAA ACAGAACACT TTTCCTTCCA TGTGGTCTGA ATGTTGGCAC CAGCCCGGGC 840
AGGGGCATCT CATTGTGGCA GTACTGCTGT GCAACCCAGC TGCAAGGATG GAAGGCAGAG 900
GGTGGGTGTG GGGCCTGAGG CTTCACAGTA CCTGGACCAG CAGGAAGATT CTGGGAGGTC 960
ACTGCTCTCA GAGGACAGCA AGGGACCCTG AGCTCTGCAA GCTGTGATCT GTCTGGGTTC 1020
ATGGTTTTTC TCAAATCCCA GGCTATCTGC ATGCGCTCTC AGGTGCTACC GAGCCATCCT 1080
GGGAGAGATG GATGGTCCAC TGCTTTGAGG CAGGGAGCCA TCGGGCTGGG GCCCCTTGGT 1140
GAACCTGATG CAGGTAAGAT GCTGAGGACT AAAACCATT TTTTTCGACC CAAAAAAAAA 1200
GGCAGGAAAA TGATCATCAG AAATAAATG GCAGCCAGGC ATGGGGGCTC ACGACTGTAA 1260
TCCTCGCACT TTGGGAGGCT CAGGCTAAGG GTCGCTTGAA GCTGAGAGTT CAAGACCAAC 1320
CTGGGCAACA TAGTGAGACC CCCATCTCTA CAATTTTTTT TTAATGACCA AATGTGGCGG 1380
TACATACCTG TACTACCTG CGGTTCCAGC TACTCAAGAG GCTGAGGCAG GAGGACTGCT 1440
TGAGCCCAGG AGATTACGGC TGCAGTGAG TACGATCAAG CCACTGCACT CCAGCCTGGG 1500
CGACAGAGCA AGATCGTTTC TCTAAAATT
```

25 Seq ID NO: 122 Protein sequence:  
Protein Accession #: NP\_060776

30

```
1 11 21 31 41 51
| | | | |
METSASSSQP QDNSQVHRET EDVDYGETDF HKQDGKAGLF SQEQYERNKS SSSSFSSSSS 60
SSSSSSSSSS GPGHGEPPDL KDELQLYGDA PGEVVPSEGS GLRRRGSDPA SGEVEASQLR 120
RLNIKKDDEF FHFVLLCFAI GALLVCYHYY ADWFMSLVG LFTFASLETV GIYFGLVYRI 180
HSLVQGFIPL FQKFRLTGFR KTD
```

35 Seq ID NO: 123 DNA sequence  
Nucleic Acid Accession #: BC022542  
Coding sequence: 243..896

40  
45  
50  
55  
60  
65  
70

```
1 11 21 31 41 51
| | | | |
ACTTGGTCCC AGCCGATAAA TCTGGGGCAG CGCGCGGTAG GAGCTGCGGG CGGCCAGGCC 60
CCTTCCTGCG TCCGCACCTG GCCCCGCGCG CCCCTCTCGG GCGTCCGGCT TCCGGCGTCC 120
TGGCGGCTCG GGTGGCGGCG GTTCGGGCGG CCGCCTGGCT GCTCCTCGGG GCGGCGACGG 180
GGCTCACGCG CGGGCCCGCC ACGGCCTTCA CCGCCGCGCG CTCTGACGCC GGCATAAGGG 240
CCATGTGTTT TGAAATTATT TTGAGGCAAG AAGTTTTGAA AGATGGTTTC CACAGAGACC 300
TTTAAATCAA AGTGAAGTTT GGGGAAAGCA TTGAGGACTT GCACACGTGC CGTCTCTTAA 360
TTAAACAGGA CATTCTGCA GGACTTTATG TGGATCCGTA TGAGTTGGCT TCATTACGAG 420
AGAGAAACAT AACAGAGGCA GTGATGGTTT CAGAAAATTT TGATATAGAG GCCCCTAACT 480
ATTTGTCCAA GGAGTCTGAA GTTCTCATTT ATGCCAGACG AGATTCACAG TGCATTGACT 540
GTTTTCAAGC CTTTTTGCC TGTCACTGCC GCTATCATCG GCCGCACAGT GAAGATGGAG 600
AAGCCTCGAT TGTGGTCAAT AACCCAGATT TGTGTGTTT TTGTGACCAA GAGTTCCCGA 660
TTTTGAAATG CTGGGCTCAC TCAGAAAGTG CAGCCCCTTG TGCTTTGGAT AATGAGGATA 720
TATGCCAATG GAACAAGATG AAGTATAAAT CAGTATATAA GAATGTGATT CTACAAGTTC 780
CAGTGGGACT GACTGTACAT ACCTCTCTAG TATGTTCTGT GACTCTGCTC ATTACAATCC 840
TGTGCTCTAC ATTGATCCTT GTAGCAGTTT TCAAATATGG CCATTTTTC CTATAAGTTT 900
TATGTAGTTA AATGCTTCTT AGAAACCTAA ATAAGATCTA TTAATTTCTG ACGAGAGGTG 960
TTCTTCTAGA ATTAATTACT TTTATCTTTT GTCTTCATTT GTGGCCAAAA TTATGTTTAC 1020
TAGAGGAAAT TTGGGATCAT TCTCAGCTAA TTCCAAAATG TAGTGCTCTA TTGCATGGAT 1080
CCTTGGAAT CACTAAGCAT CAGATGCCAT AAGGGGAAAC TTAATCTGCT TAAATTAATG 1140
TTTATTTTGT GAGAAGTGAC TTTATCTTCA TTTGGGGTAG AAAAATTATT TCTTTATGTA 1200
GTAGAGACAA ATTATTCTCA TTTTGCAAGT ACTTTCAATT TAAGCTACAA ATTGAGAAAA 1260
CCGTATATAA TAAGAATAAA ATAGGCCAGG CACAGTGGCT CACACCTGTA ATCCCAGCAC 1320
TTTGGGAGGC CGAGGTGGGC GGATCACCAG AGGTCAAGAG TTGAGACCA GCTTGGTGAA 1380
ACCCTGTCTC TACTAAAAAT ACAAAGTTA GCTGGGCTG GTGGTGGGCA TCTGTAGTCC 1440
CAGCTAATTG GAAGGGTGAG GCGGGAGGAT CGCTTGAACC TGGGAGGCGG AGGTTCCAGA 1500
GAGCCAAGAT CGCACCCTG CACTACAGCC TGGGCGACAG AACGAGACCC TGTCTCCAAA 1560
GGAAAAACAA AAAAGAAGAA TAAATAATT TGGATGAAAA TCATGTTTAT TTAATAGTA 1620
ATGTCATGAG ACTATTAAAG ATGTGCCAGA GTTTCAATGA AAATCATTA AGTAGGACAG 1680
CTAAGAAATT AATATTAATA TAAAAATTAT TGATAATCTT AAATTATTGA TTATTCCTTA 1740
ACGCACTCCA TTCTCCTTTT ACATTTTATC ATGTTTCTTT TGAATATATG AATTGGCAAA 1800
GGACTTGATG AAAGTGAGTA CTAAGATTG GTACAGAGTA TGTCAGGAAG ACAACTCAGA 1860
TTGCCATTTT AAATAAAGTT GTACATGAAC AAAAAAAAAA AAAAAA
```

75 Seq ID NO: 124 Protein sequence:  
Protein Accession #: AAH22542

80

```
1 11 21 31 41 51
| | | | |
MCSEILRLQE VLKDGFRDL LIKVKFGESI EDLHTCRLLI KQDIPAGLYV DPYELASLRE 60
RNITEAVMVS ENFDIEAPNY LSKESEVLIY ARRDSQCIDC FQAFPLPVHCR YHRPHSEDEGE 120
ASIVVNNPDL LMFCDQAGSR RMIRFRFDSF DKTIEFPILK CWAHSEVAAP CALENEDICQ 180
WNNMKYKSVY KNVILQVPVG LTVHTSLVCS VTLLITILCS KKKKK
```

85 Seq ID NO: 125 DNA sequence  
Nucleic Acid Accession #: NM\_004994.1  
Coding sequence: 20..2143



1 11 21 31 41 51  
 | | | | | |  
 AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCCTG GTCCTGGTGC TCCTGGTGCT 60  
 GGGCTGCTGC TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTTGTGCTCT TCCCTGGAGA 120  
 5 CCTGAGAACC AATCTCACCG ACAGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGTTA 180  
 CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTTCT 240  
 CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300  
 GCGAACCCCA CGGTGCGGGG TCCCAGACCT GGGCAGATTG CAAACCTTTG AGGGCGACCT 360  
 CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTGCCGCG 420  
 10 GCGGGTGATT GACGACGCCT TTGCCCAGCG CTTCGCACTG TGGAGCGCGG TGACGCCGCT 480  
 CACCTTCACT CGCGTGTACA GCCGGGACGC AGACATCGTC ATCCAGTTTG GTGTGCGCGA 540  
 GCACGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACG CCTTTCCTCC 600  
 TGGCCCCGGC ATTCAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCTGGGCAA 660  
 GGGCGTCTGT GTTCCAACCT GGTTCGAAA CGCAGATGGC GCGGCCTGCC ACTTCCCCTT 720  
 15 CATCTTCGAG GGCCGCTCCT ACTCTGCCTG CACCACCGAC GGTCGCTCCG ACGGCTTGCC 780  
 CTGGTGCACT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGCTTCT GCGCCAGCGA 840  
 GAGACTCTAC ACCCGGGACG GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT 900  
 CCAAGGCCAA TCCTACTCCG CCTGCACCAC GGACGGTCCG TCCGACGGCT ACCGCTGGTG 960  
 CGCCACCACC GCCAAGTACG ACCGGGACAA GCTCTTCGGC TTCTGCCCGA CCGAGCTGA 1020  
 20 CTCGACGGTG ATGGGGGGCA ACTCGGCGGG GCGAGCTGTG GTCTTCCCCT TCACCTTCTT 1080  
 GGGTAAGGAG TACTCGACCT GTACCAGCGA GGGCCGCGGA GATGGGCGCC TCTGGTGCGC 1140  
 TACCACCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200  
 TTTGTTCTCT GTGGCGGCGC ATGAGTTCGG CCACGCGCTG GGCTTAGATC ATTCCTCAGT 1260  
 25 GCCGGAGGCG CTCATGTACC CTATGTACCG CTTCACTGAG GGGCCCCCTT TGCATAAGGA 1320  
 CGACGTGAAT GGCATCCGGC ACCTCTATGG TCCTCGCCCT GAACCTGAGC CACGGCCTCC 1380  
 AACCACCACC ACACCGCAGC CCACGGCTCC CCGACGGTTC TGCCCCACCG GACCCCCCAC 1440  
 TGTCCACCCC TCAGAGCGCC CCACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC 1500  
 AGGTCCCCCC ACTGCTGGCC CTTCTACGGC CACTACTGTG CTTTGTAGTC CGGTGGACGA 1560  
 30 TGCTTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAAACAGC TGTATTTGTT 1620  
 CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGGG AGCCGGCCGC AGGGCCCCCT 1680  
 CCTTATCGCC GACAAGTGGC CCGCGCTGCC CCGCAAGCTG GACTCGGTCT TTGAGGAGCC 1740  
 GCTCTCCAAG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC 1800  
 GGTGCTGGGC CCGAGGCGTC TGGACAAGCT GGGCCTGGGA GCCGACGTGG CCCAGGTGAC 1860  
 35 CGGGGCCCTC CGGAGTGGCA GGGGGAAGAT GCTGCTGTTT AGCGGGCGGC GCCTCTGGAG 1920  
 GTTCGACGTG AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGATGTT 1980  
 CCCCAGGGTG CCTTTGGACA CGCACGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG 2040  
 CCAGGACCGC TTCTACTGGC GCGTGAGTTC CCGGAGTGAG TTGAACCAGG TGGACCAAGT 2100  
 GGGCTACGTG ACCTATGACA TCTGTGAGTG CCCTGAGGAC TAGGGCTCCC GTCTGCTTT 2160  
 40 GCAGTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCCGATA 2220  
 CAAACTGGTA TTCTGTTCTG GAGGAAAGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC 2280  
 TCACCTTTGT TTTTGTGTTG AGTGTCTTA ATAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 126 Protein sequence:  
 Protein Accession #: NP\_004985.1

1 11 21 31 41 51  
 | | | | | |  
 MSLWQPLVLV LLVLGCCFAA PRQRQSTLVL FPGDLRTNLT DRQLAEEYLY RYGYTRVAEM 60  
 RGSKSLGPA LLLLQKQLSL PETGELDSAT LKAMRTPRCG VPDLGRFQTF EGDLEKWHHHN 120  
 50 ITYWIQNYSE DLPRVIDDA FARAFALWSA VTPLTFTRVY SRDADIVIQF GVAEHGDGYP 180  
 FDGKDGLLAH AFPPGPGIQG DAHFDDELDW SLGKGVVPT RFGNADGAAC HFPFIFEGRS 240  
 YSACTTDGRS DGLPWCSTTA NYDTRDRFGF CPSERLYTRD GNADGKPCQF PFIFQGQSYS 300  
 ACTTDGRSDG YRWCATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTFLGKEYST 360  
 55 CTSEGRGDGR LWCATTSNFD SDKKWGFCDP QGYSLFLVAA HEFGHALGLD HSSVPEALMY 420  
 PMYRFTEGPP LHKDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAPPTVCPT GPPTVHPSER 480  
 PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFDKDKYW 540  
 RFSEGRGSRP QGPFLIADKW PALPRKLDV FFEPLSKKLF FFSGRQVWVY TGASVLGPRR 600  
 LDKLGLGADV AQVTGALRSR RGKMLLFSGR RLWRFVDVKAQ MVDPRSASEV DRMFPGVPLD 660  
 60 THDVFYREK AYFCQDRFYW RVSSRSELNQ VDQVGYVTYD ILQCPED

Seq ID NO: 127 DNA sequence  
 Nucleic Acid Accession #: NM\_004181  
 Coding sequence: 32-670

1 11 21 31 41 51  
 | | | | | |  
 GCAGAAATAG CCTAGGGAGA TCAACCCCGA GATGCTGAAC AAAGTGCTGT CCCGGCTGGG 60  
 GGTGCGCCGC CAGTGGCGCT TCGTGGACGT GCTGGGGCTG GAAGAGGAGT CTCTGGGCTC 120  
 70 GGTGCCAGCG CCTGCCTGCG CGCTGCTGCT GCTGTTTCCC CTCACGGCCC AGCATGAGAA 180  
 CTTACAGAAA AAGCAGATTG AAGAGCTGAA GGGACAAGAA GTTAGTCCTA AAGTGTACTT 240  
 CATGAAGCAG ACCATTGGGA ATTCCTGTGG CACAATCGGA CTTATTACG CAGTGGCCAA 300  
 TAATCAAGAC AAAGTGGGAT TTGAGGATGG ATCAGTTCTG AAACAGTTTC TTTCTGAAAC 360  
 75 AGAGAAAATG TCCCCTGAAG ACAGAGCAAA ATGCTTTGAA AAGAATGAGG CCATACAGGC 420  
 AGCCCATGAT GCCGTGGCAC AGGAAGGCCA ATGTCGGGTA GATGACAAGG TGAATTTCCA 480  
 TTTTATTCTG TTTAACAACG TGGATGGCCA CCTCTATGAA CTTGATGGAC GAATGCCTTT 540  
 TCCGGTGAAC CATGGCGCCA GTTCAGGGA CACCTGCTG AAGGACGCTG CCAAGGTGTG 600  
 CAGAGAATTC ACCGAGCGTG AGCAAGGAGA AGTCCGCTTC TCTGCCGTGG CTCTCTGCAA 660  
 GGCAGCCTAA TGCTCTGTGG GAGGGACTTT GCTGATTTCC CCTCTTCCCT TCAACATGAA 720  
 80 AATATATACC CCCCATGCAG TCTAAAATGC TTCAGTACTT GTGAAACACA GCTGTTCTTC 780  
 TGTCTGCAG ACACGCCTTC CCCTCAGCCA CACCCAGGCA CTTAAGCACA AGCAGAGTGC 840  
 ACAGCTGTCC ACTGGGCCAT TGTGGTGTGA GCTTCAGATG GTGAAGCATT CTCCCAGTG 900  
 TATGTCTTGT ATCCGATATC TAACGCTTTA AATGGCTACT TTGGTTTCTG TCTGTAAAGT 960  
 AAGACCTTGG ATGTGTTTAT GTTGTCTTAA AGAATAAATT TTGCTGATAG TAGC

Seq ID NO: 128 Protein sequence:  
 Protein Accession #: NP\_004172



1	11	21	31	41	51	
MLNKVLSRLG	VAGQWRFDV	LGLEESLGS	VPAPACALLL	LFPLTAQHEN	FRKKQIEELK	60
GQEVSPKVYF	MKQTIGNSCG	TIGLIHAVAN	NQDKLGFEDG	SVLKQFLSET	EKMSPEDRAK	120
CFEKNEAIQA	AHDAVAQEGQ	CRVDDKVNFI	FILENNVDGH	LYELDGRMPF	PVNHGASSED	180
TLLKDAAKVC	REFTEREQGE	VRFSVALCK	AA			

Seq ID NO: 129 DNA sequence  
Nucleic Acid Accession #: NM\_000213  
Coding sequence: 127-5385

1	11	21	31	41	51	
CGCCCCGCGC	CTGCAGCCCC	ATCTCCTAGC	GGCAGCCCCG	GCGCGGAGGG	AGCGAGTCCG	60
CCCCGAGGTA	GGTCCAGGAC	GGGCGCACAG	CAGCAGCCGA	GGCTGGCCGG	GAGAGGGAGG	120
AAGAGGATGG	CAGGGCCACG	CCCCAGCCCA	TGGGCCAGGC	TGCTCCTGGC	AGCCTTGATC	180
AGCGTCAGCC	TCTCTGGGAC	CTTGGCAAAC	CGCTGCAAGA	AGGCCCCAGT	GAAGAGCTGC	240
ACGGAGTGTG	TCCGTGTGGA	TAAGGACTGC	GCCTACTGCA	CAGACGAGAT	GTTCAAGGAC	300
CGGCGCTGCA	ACACCCAGGC	GGAGCTGTGT	GCCGCGGGCT	GCCAGCGGGA	GAGCATCGTG	360
GTCATGGAGA	GCAGCTTCCA	AATCACAGAG	GAGACCCAGA	TTGACACCAC	CCTGCGGCGC	420
AGCCAGATGT	CCCCCAAGG	CCTGCGGGTC	CGTCTGCGGC	CCGGTGAGGA	GCGGCATTTT	480
GAGCTGGAGG	TGTTTGAGCC	ACTGGAGAGC	CCCGTGGACC	TGTACATCCT	CATGGACTTC	540
TCCAACCTCA	TGTCCGATGA	TCTGGACAAC	CTCAAGAAGA	TGGGGCAGAA	CCTGGCTCGG	600
GTCTTGAGCC	AGCTCACCAG	CGACTACACT	ATTGGATTGT	GCAAGTTTGT	GGACAAAGTC	660
AGCGTCCCGC	AGACGGACAT	GAGGCCTGAG	AAGCTGAAGG	AGCCCTGGCC	CAACAGTGAC	720
CCCCCTTCT	CCTTCAAGAA	CGTCATCAGC	CTGACAGAAG	ATGTGGATGA	GTTCCGGAAT	780
AAACTGCAGG	GAGAGCGGAT	CTCAGGCAAC	CTGGATGCTC	CTGAGGGCGG	CTTCGATGCC	840
ATCCTGCAGA	CAGCTGTGTG	CACGAGGGAC	ATTGGCTGGC	GCCCGGACAG	CACCCACCTG	900
CTGGTCTTCT	CCACCGAGTC	AGCCTTCCAC	TATGAGGCTG	ATGGCGCCAA	CGTGTCTGGT	960
GGCATCATGA	GCGCAACGA	TGAACGGTGC	CACCTGGACA	CCACGGGCAC	CTACACCCAG	1020
TACAGGACAC	AGGACTACCC	GTCGGTGCCC	ACCCTGGTGC	GCCTGCTCGC	CAAGCACAAAC	1080
ATCATCCCCA	TCTTTGCTGT	CACCAACTAC	TCCTATAGCT	ACTACGAGAA	GCTTCACACC	1140
TATTTCCCTG	TCTCCTCACT	GGGGGTGTGT	CAGGAGGACT	CGTCCAACAT	CGTGGAGCTG	1200
CTGGAGGAGG	CCTTCAATCG	GATCCGCTCC	AACCTGGACA	TCCGGGCCCT	AGACAGCCCC	1260
CGAGGCCCTC	GGACAGAGGT	CACCTCCAAG	ATGTTCCAGA	AGACGAGGAC	TGGGTCTTTT	1320
CACATCCGGC	GGGGGGAAGT	GGGTATATAC	CAGGTGCAGC	TGCGGGCCCT	TGAGCACGTG	1380
GATGGGACGC	ACGTGTGCCA	GCTGCGGAG	GACCAGAAGG	GCAACATCCA	TCTGAAACCT	1440
TCCTTCTCCG	ACGGCCTCAA	GATGGACGCG	GGCATCATCT	GTGATGTGTG	CACCTGCGAG	1500
CTGCAAAAAG	AGGTGCGGTC	AGCTCGCTGC	AGCTTCAACG	GAGACTTCGT	GTGCGGACAG	1560
TGTGTGTGCA	GCGAGGGCTG	GAGTGGCCAG	ACCTGCAACT	GCTCCACCGG	CTCTCTGAGT	1620
GACATTCAGC	CCTGCCTGCG	GGAGGGCGAG	GACAAGCCGT	GCTCCGGCCG	TGGGGAGTGC	1680
CAGTGCGGGC	ACTGTGTGTG	CTACGGCGAA	GGCCGCTACG	AGGGTCAGTT	CTGCGAGTAT	1740
GACAACCTTC	AGTGTCCCCG	ACTTCCGGG	TTCTCTGCA	ATGACCGAGG	ACGCTGCTCC	1800
ATGGGCCAGT	GTGTGTGTGA	GCCTGGTTGG	ACAGGCCCAA	GCTGTGACTG	TCCCCTCAGC	1860
AATGCCACCT	GCATCGACAG	CAATGGGGGC	ATCTGTAATG	GACGTGGCCA	CTGTGAGTGT	1920
GGCCGCTGCC	ACTGCCACCA	GCAGTCGCTC	TACACGGACA	CCATCTGCGA	GATCAACTAC	1980
TCGGCGATCC	ACCCGGGCCCT	CTGCGAGGAC	CTACGCTCCT	GCGTGCACTG	CCAGGCGTGG	2040
GGCACC GGCG	AGAAGAAGGG	GCGCAGTGT	GAGGAATGCA	ACTTCAAGGT	CAAGATGGTG	2100
GACGAGCTTA	AGAGAGCCGA	GGAGGTGGTG	GTGCGCTGCT	CCTTCCGGGA	CGAGGATGAC	2160
GACTGCACCT	ACAGCTACAC	CATGGAAGGT	GACGGCGCCC	CTGGGCCCAA	CAGCACTGTC	2220
CTGGTGCACA	AGAAGAAGGA	CTGCCCTCCG	GGCTCCTTCT	GGTGGCTCAT	CCCCCTGCTC	2280
CTCCTCCTCC	TGCCGCTCCT	GGCCTGCTA	CTGCTGCTAT	GCTGGAAGTA	CTGTGCCTGC	2340
TGCAAGGCCT	GCTTGGCACT	TCTCCCGTGC	TGCAACCGAG	GTCACATGGT	GGGCTTTAAG	2400
GAAGACCACT	ACATGCTGCG	GGAGAACCCT	ATGGCCTCTG	ACCACTTGGA	CACGCCCATG	2460
CTGCGCAGCG	GGAACTCAA	GGGCGGTGAC	GTGGTCCGCT	GGAAGGTAC	CAACAACATG	2520
CAGCGGCCTG	GCTTTGCCAC	TCATGCCGCC	AGCATCAACC	CCACAGAGCT	GGTGCCTTAC	2580
GGGCTGTCTC	TGCGCCTGGC	CCGCCTTTGC	ACCGAGAACC	TGCTGAAGCC	TGACACTCGG	2640
GAGTGCGCCC	AGCTGCGCCA	GGAGGTGGAG	GAGAACCCTGA	ACGAGGTCTA	CAGGCAGATC	2700
TCCGGTGTAC	ACAAGCTCCA	GCAGACCAAG	TTCCGGCAGC	AGCCCAATGC	CGGGAAAAAG	2760
CAAGACCACA	CCATTGTGGA	CACAGTGCTG	ATGGCGCCCC	GCTCGGCCAA	GCCGGCCCTG	2820
CTGAAGCTTA	CAGAGAAGCA	GGTGAACAG	AGGGCCTTCC	ACGACCTCAA	GGTGGCCCCC	2880
GGCTACTACA	CCCTCACTGC	AGACCGGAC	GCCCGGGGCA	TGGTGGAGTT	CCAGGAGGGC	2940
GTGGAGCTGG	TGGACGTACG	GGTGCCCTC	TTTATCCGGC	CTGAGGATGA	CGACGAGAAG	3000
CAGCTGCTGG	TGGAGGCCAT	CGACGTGCCC	GCAGGCACTG	CCACCCTCGG	CCGCCGCTG	3060
GTAAACATCA	CCATCATCAA	GGAGCAAGCC	AGAGACGTGG	TGTCTTTTGA	GCAGCCTGAG	3120
TTCTCGGTCA	GCCGCGGGGA	CCAGGTGGCC	CGCATCCCTG	TCATCCGGCG	TGTCCTGGAC	3180
GGCGGGGAAGT	CCCAGGTCTC	CTACCGCACA	CAGGATGGCA	CCGCGCAGGG	CAACCGGGAC	3240
TACATCCCCG	TGGAGGGTGA	CAGCTGTTC	GAGCCTGGGG	AGGCCTGGAA	AGAGCTGCAG	3300
GTGAAGCTCC	TGGAGCTGCA	AGAAGTTGAC	TCCCTCCTGC	GGGGCCGCCA	GGTCCGCCGT	3360
TTCCACGTCC	AGCTCAGCAA	CCCTAAGTTT	GGGGCCACC	TGGGCCAGCC	CCACTCCACC	3420
ACCATCATCA	TCAGGGACCC	AGATGAACTG	GACCGGAGCT	TCACGAGTCA	GATGTTGTCA	3480
TCACAGCCAC	CCCCTCACGG	CGACCTGGGC	GCCCCGCAGA	ACCCCAATGC	TAAGGCCGCT	3540
GGGTCCAGGA	AGATCCATTT	CAACTGGCTG	CCCCCTTCTG	GCAAGCCAAT	GGGGTACAGG	3600
GTAAAGTACT	GGATTGAGG	TGACTCCGAA	TCCGAAGCCC	ACCTGCTCGA	CAGCAAGGTG	3660
CCCTCAGTGG	AGCTCACCAA	CCTGTACCCG	TATTGCGACT	ATGAGATGAA	GGTGTGCGCC	3720
TACGGGGCTC	AGGGCGAGGG	ACCCTACAGC	TCCCTGGTGT	CCTGCCGCAC	CCACCAGGAA	3780
GTGCCCAGCG	AGCCAGGGCG	TCTGGCCTTC	AATGTCGTCT	CCTCCACGGT	GACCCAGCTG	3840
AGCTGGGCTG	AGCCGCTGGA	GACCAACGGT	GAGATCACAG	CCTACGAGGT	CTGCTATGGC	3900
CTGGTCAACG	ATGACAACCG	ACCTATTGGG	CCCATGAAGA	AAGTGCTGGT	TGACAACCTT	3960
AAGAACC GGA	TGCTGCTTAT	TGAGAACCTT	CGGGAGTCCC	AGCCCTACCG	CTACACGGTG	4020
AAGGCGCGCA	ACGGGGCCCG	CTGGGGCCCT	GAGCGGGAGG	CCATCATCAA	CCTGGCCACC	4080
CAGCCCAAGA	GGCCCATGTC	CATCCCCATC	ATCCCTGACA	TCCCTATCGT	GGACGCCACG	4140
AGCGGGGAGG	ACTACGACAG	CTTCCTTATG	TACAGCGATG	ACGTTCTACG	CTCTCCATCG	4200
GGCAGCCAGA	GGCCAGCGT	CTCCGATGAC	ACTGAGCACC	TGGTGAATGG	CCGGATGGAC	4260
TTTGCCTTCC	CGGGCAGCAC	CAACTCCCTG	CACAGGATGA	CCACGACCAG	TGCTGCTGCC	4320
TATGGCACCC	ACCTGAGCCC	ACACGTGCCC	CACCGCGTGC	TAAGCACATC	CTCCACCCTC	4380



ACACGGGACT ACAACTCACT GACCCGCTCA GAACACTCAC ACTCGACCAC ACTGCCGAGG 4440  
 GACTACTCCA CCCTCACCTC CGTCTCCTCC CACGACTCTC GCCTGACTGC TGGTGTGCCC 4500  
 GACACGCCCA CCCGCTGGT GTTCTCTGCC CTGGGGCCCA CATCTCTCAG AGTGAGCTGG 4560  
 CAGGAGCCGC GGTGCGAGCG GCCGCTGCAG GGCTACAGTG TGGAGTACCA GCTGCTGAAC 4620  
 GCGGCTGAGC TGCATCGGCT CAACATCCCC AACCTGCCCC AGACCTCGGT GGTGGTGGAA 4680  
 GACCTCCTGC CCAACCACTC CTACGTGTTT CGCGTGCAGG CCCAGAGCCA GGAAGGCTGG 4740  
 GGCCGAGAGC GTGAGGGTGT CATCACCATT GAATCCCAGG TGACCCGCA GAGCCCACTG 4800  
 TGTCCCCTGC CAGGCTCCGC CTTCACTTTG AGCACTCCCA GTGCCCCAGG CCCGCTGGTG 4860  
 TTCCTGCCCC TGAGCCCAGA CTCGCTGCAG CTGAGCTGGG AGCGGCCACG GAGGCCCAAT 4920  
 GGGGATATCG TCGGCTACCT GGTGACCTGT GAGATGGCCC AAGGAGGAGG GCCAGCCACC 4980  
 GCATTCGCGG TGGATGGAGA CAGCCCCGAG AGCCGGCTGA CCGTGCCGGG CCTCAGCGAG 5040  
 AACGTGCCCT ACAAGTTCAA GGTGCAGGCC AGGACCACTG AGGGCTTCGG GCCAGAGCGC 5100  
 GAGGGCATCA TCACCATAGA GTCCCAGGAT GGAGGACCTT TCCCGCAGCT GGGCAGCCGT 5160  
 GCGGGGCTCT TCCAGCACCC GCTGCAAAGC GAGTACAGCA GCATCACAC CACCCACACC 5220  
 AGCGCCACCG AGCCCTTCCT AGTGGATGGG CCGACCCTGG GGGCCCAGCA CCTGGAGGCA 5280  
 GGCGGCTCCC TCACCCGGCA TGTGACCCAG GAGTTTGTGA GCCGGACACT GACCACCAGC 5340  
 GGAACCCCTA GCACCCACAT GGACCAACAG TTCTTCCAAA CTTGACCGCA CCCTGCCCA 5400  
 CCCCCGCCAT GTCCCACTAG GCGTCTCTCC GACTCTCTC CCGGAGCCTC CTCAGCTACT 5460  
 CCATCCTTGC ACCCCTGGG GCCCAGCCCA CCGCATGCA CAGAGCAGG GCTAGGTGTC 5520  
 TCCTGGGAGG CATGAAGGGG GCAAGGTCCG TCCTCTGTGG GCCCAAACCT ATTTGTAACC 5580  
 AAAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTCTG CACTTAATAA ATGGTTTTCG 5640  
 ACTG

Seq ID NO: 130 Protein sequence:  
 Protein Accession #: NP\_000204

1 11 21 31 41 51  
 | | | | |  
 MAGPRPSPA RLLLAALISV SLSGTLANRC KKAPVKSCTE CVRVDKDCAY CTDEMFRDRR 60  
 CNTQAEELAA GCQRESIVVM ESSFQITEET QIDTTLRRSQ MSPQGLRVRL RPGEERHFEL 120  
 EVFEPLESPV DLYILMDFSN SMSDDLNLK KMGQNLARVL SOLTSDYTIG FGKFDVKVSV 180  
 PQTDMRPEKL KEPWPNSDPP FSFKNVISLT EDVDEFNRKL QGERISGNLD APEGGFDAIL 240  
 QTAVCTRDIG WRPDSTHLLV FSTESAFHYE ADGANVLAGE MSRNDERCHL DTTGTYTQYR 300  
 TQDYPSPVPTL VRLAKHNII PIFAVTNSY SYEKLHTYF PVSSLGVLQE DSSNIVELLE 360  
 EAFNRIRSNL DIRALDSPRG LRTEVTSKMF QKTRTGSFHI RRGEVGIYQV QLRALHVDG 420  
 THVCQLPEDQ KGNHLKPSF SDGLKMDAGI ICDVCTCELQ KEVRSARCSF NGDFVCGQCV 480  
 CSEGWSGQTC NCSTGSLSDI QPCLREGEDK PCSGRGECQC GHCVCYGEGR YEGQFCEYDN 540  
 FQCPRTSGFL CNDRGRCMSG QCVCPEPGWTG PSCDCPLSNA TCIDSNGGIC NGRGHCECGR 600  
 CHCHQQSLYT DTICEINYSI IHPGLCEDLR SCVQCQAWGT GEKKGRTCEE CNFKVKMVDE 660  
 LKRAEEVVVR CSFRDEDDDC TYSYTMEDGD APGPNSTVLV HKKKDCPPGS FWWLIPLLL 720  
 LLPLLALLLL LCWKYCACCK ACLALLPCCN RGHMVGFKEG HYMLRENLMA SPHLDTPLMR 780  
 SGNLKG RDVW RWKVTNNMQR PGFATHAASI NPTELVPYGL SLRLARLCTE NLLKPDTRER 840  
 AQLRQEVEEN LNEVYRQISG VHKLQQTQKR QQPNAKKQD HTIVDTVLMA PRSAKPALK 900  
 LTEKQVEQRA FHDLLKVPAGY YTLTADQDAR GMVEFQEGVE LVDVRVPLFI RPEDDDEKQL 960  
 LVEAIDVPAG TATLGRRLLV ITIIKEQARD VVSFEQPEFS VSRGDQVARI PVIRRVLDGG 1020  
 KSQVSYRTQD GTAQGNRDYI PVEGELLFQP GEAWKELQVK LLELQEVDSL LRGRQVRRFH 1080  
 VQLSNPKFGA HLGQPHSTTI IIRDPDELDR SFTSQMLSSQ PPPHGDLAGP QNPNAKAAGS 1140  
 RKIHFNLWLP SGKPMGYRVK YWIQGDSESE AHLDSKVPS VELTNLYPYC DYEMKVCAYG 1200  
 AQGEGPYSSL VSCRTHQEVN SEPGRFAFNV VSSTVTQLSW AEPAETNGEI TAYEVCYGLV 1260  
 NDDNRPIGPM KKVLDVNPKN RMLLIENLRE SQPYRYTVKA RAGAGWGPER EAIINLATQP 1320  
 KRPMISIPPI DIPIVDAQSG EDYDSFLMYS DDVLRSPSGS QRPSVSDDE HLVNRMDFDA 1380  
 FPGSTNSLHR MTTTSAAYG THLSPHVPHR VLSTSTLTR DYNLSLTRSEH SHSTTLPRDY 1440  
 STLTSSVSHD SRLTAGVPDT PTRLVFSALG PTSLRVSWQE PRCERPLQGY SVEYQLLNGG 1500  
 ELHRLNIPNP AQTSSVVEDL LPNHSYVFRV RAQSQEGWGR EREGVITIES QVHPQSPICP 1560  
 LPGSAFTLST PSAPGPLVFT ALSPDSLQLS WERPRRPNGD IVGYLVTCEM AQGGGPATAF 1620  
 RVDGDSPESE LTVPGLSENV PYKFKVQART TEGFGPEREG IITIESQDGG PFPQLGSRAG 1680  
 LFQHPQLQSEY SSITTTHTSA TEPFLVDGPT LGAQHLEAGG SLTRHVTQEF VSRTLTTSGT 1740  
 LSTHMDQQFF QT

Seq ID NO: 131 DNA sequence  
 Nucleic Acid Accession #: BC004372  
 Coding sequence: 132..2231

1 11 21 31 41 51  
 | | | | |  
 CCTCGTGCCG CGGACCCAG CCTCTGCCAG GTTCGGTCCG CCATCCTCGT CCCGTCCTCC 60  
 GCCGGCCCCCT GCCCGCGCC CAGGGATCCT CCAGCTCCTT TCGCCCGCGC CCTCCGTTCCG 120  
 CTCGGACAC CATGGACAAG TTTTGGTGGC ACGCAGCCTG GGGACTCTGC CTCGTGCCGC 180  
 TGAGCCTGGC GCAGATCGAT TTGAATATAA CTTGCCGCTT TGAGGTGTA TTCCACGTGG 240  
 AGAAAAATGG TCGCTACAGC ATCTCTCGGA CGGAGGCCGC TGACCTCTGC AAGGCTTTCA 300  
 ATAGCACCTT GCCACAATG GCCCAGATGG AGAAAGCTCT GAGCATCGGA TTTGAGACCT 360  
 GCAGGTATGG GTTCATAGAA GGGCATGTGG TGATTCCTCC GATCCACCCC AACTCCATCT 420  
 GTGCAGCAAA CAACACAGGG GTGTACATCC TCACATCCAA CACCTCCAG TATGACACAT 480  
 ATTGCTTCAA TGCTTCAGCT CCACCTGAAG AAGATTGTAC ATCAGTCACA GACCTGCCCA 540  
 ATGCTTTTGA TGGACCAATT ACCATAACTA TTGTTAACC TGATGGCACC CGCTATGTCC 600  
 AGAAAGGAGA ATACAGAAGC AATCCTGAAG ACATCTACCC CAGCAACCTT ACTGATGATG 660  
 ACGTGAGCAG CGGCTCCTCC AGTGAAAGGA GCAGCACTTC AGGAGGTTAC ATCTTTTACA 720  
 CCTTTTCTAC TGTACACCCC ATCCAGACG AAGACAGTCC CTGGATCACC GACAGCACAG 780  
 ACAGAAATCCC TGCTACCAGT ACGTCTTCAA ATACCATCTC AGCAGGCTGG GAGCCAAATG 840  
 AAGAAATGA AGATGAAAGA GACAGACACC TCAGTTTTTC TGGATCAGGC ATTGATGATG 900  
 ATGAAGATT TATCTCCAGC ACCATTTCAA CCACACCAG GGCTTTTGAC CACACAAAAC 960  
 AGAACCAGGA CTGGACCCAG TGAACCCAA GCCATTCAA TCCGGAAGTG CTACTTCAGA 1020  
 CAACCACAAG GATGACTGAT GTAGACAGAA ATGGCACCAC TGCTTATGAA GGAACTGGA 1080  
 ACCCAGAAGC ACACCTCCC CTCATTACC ATGAGCATCA TGAGGAAGAA GAGACCCAC 1140  
 ATTCTACAAG CACAATCCAG GCAACTCCTA GTAGTACAAC GGAAGAAACA GCTACCCAGA 1200  
 AGGAACAGTG GTTTGGCAAC AGATGGCATG AGGGATATCG CCAAACACCC AGAGAAGACT 1260



5  
10  
15  
20

CCCATTCGAC	AACAGGGACA	GCTGCAGCCT	CAGCTCATAC	CAGCCATCCA	ATGCAAGGAA	1320
GGACAACACC	AAGCCCAGAG	GACAGTTCCT	GGACTGATTT	CTTCAACCCA	ATCTCACACC	1380
CCATGGGACG	AGGTCATCAA	GCAGGAAGAA	GGATGGATAT	GGACTCCAGT	CATAGTACAA	1440
CGCTTCAGCC	TACTGCAAAAT	CCAAACACAG	GTTTGGTGGG	AGATTTGGAC	AGGACAGGAC	1500
CTCTTTCAAT	GACAACGCAG	CAGAGTAATT	CTCAGAGCTT	CTCTACATCA	CATGAAGGCT	1560
TGGAAGAAGA	TAAAGACCAT	CCAACAACCT	CTACTCTGAC	ATCAAGCAAT	AGGAATGATG	1620
TCACAGGTGG	AAGAAGAGAC	CCAAATCATT	CTGAAGGCTC	AACTACTTTA	CTGGAAGGTT	1680
ATACCTCTCA	TTACCCACAC	ACGAAGGAAA	GCAGGACCTT	CATCCCAGTG	ACCTCAGCTA	1740
AGACTGGGTC	CTTTGGAGTT	ACTGCAGTTA	CTGTTGGAGA	TTCCAACCTC	AATGTCAATC	1800
GTTCTTATC	AGGAGACCAA	GACACATTCC	ACCCAGTGG	GGGGTCCCAT	ACCCTCATG	1860
GATCTGAATC	AGATGGACAC	TCACATGGGA	GTCAAGAAGG	TGGAGCAAAC	ACAACCTCTG	1920
GTCTTATAAG	GACACCCCAA	ATTCCAGAA	GGCTGATCAT	CTTGGCATCC	CTCTTGGCCT	1980
TGGCTTTGAT	TCTTGCAGTT	TGCATTGCAG	TCAACAGTCG	AAGAAGGTGT	GGGCAGGAAG	2040
AAAAGCTAGT	GATCAACAGT	GGCAATGGAG	CTGTGGAGGA	CAGAAAGCCA	AGTGGACTCA	2100
ACGGAGAGGC	CAGCAAGTCT	CAGGAAATGG	TGCATTGGT	GAACAAGGAG	TCGTCAGAAA	2160
CTCCAGACCA	GTTTATGACA	GCTGATGAGA	CAAGGAACCT	GCAGAATGTG	GACATGAAGA	2220
TTGGGGTGT	ACACCTACAC	CATTATCTTG	GAAAGAAACA	ACCGTTGGAA	ACATAACCAT	2280
TACAGGGAGC	TGGGACACTT	AACAGATGCA	ATGTGCTACT	GATTGTTTCA	TTGCGAATCT	2340
TTTTTAGCAT	AAAATTTTCT	ACTCTTAAAA	AAAAAAAAAA	AAAAAA		

Seq ID NO: 132 Protein sequence:  
Protein Accession #: AAH04372

25  
30  
35  
40

1	11	21	31	41	51	
MDKFWWHA	AW	GLCLVPLSL	A	QIDLNITCR	F	60
PTMAQMEK	AL	SIGFETCRY	G	FI EGHVVI	P	120
ASAPPEED	CT	SVTDLNFA	D	GPITITIV	N	180
GSSSERS	STS	GGYIFYTF	S	VHPIPED	D	240
DERDRHLS	FS	GSGIDDD	E	ISSTIST	P	300
MTDVDRNG	TT	AYEGNWN	P	EAHPPLIH	H	360
FGNRWHEG	YR	QTPRED	S	HSTGTAA	A	420
GHQAGRMD	M	DSSHSTTL	Q	PANPNTGL	V	480
KDHPFTST	LT	SSNRNDV	T	GGRRDPNH	S	540
FGVTAVTV	GD	SNSNVNR	S	LSDQDTFH	P	600
TPQIPEWLI	I	LASLLAL	AL	LAVCIAVN	R	660
SKSQEMVHL	V	NKESSE	T	PDQFMTADE	T	

Seq ID NO: 133 DNA sequence  
Nucleic Acid Accession #: NM\_002882  
Coding sequence: 150-755

45  
50  
55  
60

1	11	21	31	41	51	
CGAGGTT	CGG	GTCGTGG	GGC	GGAGGGA	AGA	60
GCGGAGG	GAA	GGAGCTA	CGA	GAGCCAG	CGA	120
AGCCGAG	CCG	CCGCGCC	CGC	CGCGCCG	CGC	180
ATGATAC	TTC	CAGTGA	AAT	ACAGAC	AGT	240
TTTCTCT	TCC	TGAGCA	AAGAA	ATTAA	AACAC	300
TGCGGG	CAAA	ACTGTT	CCGA	TTTGCT	CTG	360
GCACTGT	GTA	CGTCA	AGCTC	CTGA	AGCACA	420
GGAGGG	ACAA	GACCTG	AAG	ATCTGT	GCCA	480
AGCCCA	ACGC	AGGTAG	CGAC	CGTG	CCTGG	540
AGTGCC	CCAA	GCCAG	AGCTG	CTGG	CCATCC	600
TCAAA	ACAAA	GTTTGA	AAGAA	TGCAG	GAAAG	660
CAGGCA	AAAAA	TGATCA	TGCC	GAAAA	AGTGG	720
AGGAG	ACCAA	GGAGAT	GTCT	GAGAGA	AGC	780
TCTCTT	CTCT	TTCTTT	TTTT	TAAAA	ATTT	840
ATTCTT	TCAT	TTTTAC	AAGG	GACGTT	TATAT	
				AAAGAA	CTGA	
				ACTC		

Seq ID NO: 134 Protein sequence:  
Protein Accession #: NP\_002873

70

1	11	21	31	41	51	
MAAAKD	THED	HDTSTENT	DE	SNHDPQ	F	60
ENDLPE	WKER	GTGDV	KLLKH	KEKGA	I	120
VWNTHA	DFAD	ECPKEL	LAI	RFLNA	E	180
AEKLEA	LSVK	EETKED	AEEK	Q		

Seq ID NO: 135 DNA sequence  
Nucleic Acid Accession #: NM\_000077.2  
Coding sequence: 277-742

80  
85

1	11	21	31	41	51	
CCCAAC	CTGG	GGCGACT	TCA	GGTGTG	CCAC	60
TCCTCC	GAGC	ACTCGCT	CAC	GGCGT	CCCC	120
GGATTT	GAGG	GACAGGG	TCG	GAGGGG	GCT	180
GGGCTG	GGCTG	GTCACCA	GAG	GGTGGG	GCGG	240
GGAGAG	CAGG	CAGCGG	GCGG	CGGGG	AGCAG	300
GCCTTC	GCGCT	GACTGG	CTGG	CCACGG	CCCCG	360
GCTGCT	GAGG	GCGGGG	GCGC	TGCCCC	AACGC	420
GGTCAT	GATG	ATGGG	CAGCG	CCCGAG	TGGC	480
				GGAGCT	GTCTG	
				CTGCT	CCACG	
				GCGG	GAGCC	



CAACTGCGCC GACCCCGCCA CTCTCACCCG ACCCGTGCAC GACGCTGCCC GGGAGGGCTT 540  
 CCTGGACACG CTGGTGGTGC TGCACCGGGC CGGGGCGCGG CTGGACGTGC GCGATGCCTG 600  
 GGGCCGTCTG CCCGTGGACC TGGCTGAGGA GCTGGGCCAT CGCGATGTCG CACGGTACCT 660  
 GCGCGCGGCT GCGGGGGGCA CCAGAGGCAG TAACCATGCC CGCATAGATG CCGCGGAAGG 720  
 TCCCTCAGAC ATCCCCGATT GAAAGAACCA GAGAGGCTCT GAGAAACCTC GGGAAACTTA 780  
 GATCATCAGT CACCGAAGGT CCTACAGGGC CACAACAGCC CCCGCCACAA CCCACCCCGC 840  
 TTTTCGTAGT TTCATTTAGA AAATAGAGCT TTTAAAAATG TCCTGCCTTT TAACGTAGAT 900  
 ATATGCCTTC CCCCACTACC GTAAATGTCC ATTTATATCA TTTTATATAT ATTCTTATAA 960  
 AAATGTAAAA AAGAAAAACA CCGCTTCTGC CTTTCTACTG TGTTGGAGTT TTCTGGAGTG 1020  
 AGCACTCACG CCCTAAGCGC ACATTCATGT GGGCATTTCT TGCAGCCTC GCAGCCTCCG 1080  
 GAAGCTGTCG ACTTCATGAC AAGCATTTTG TGAAGTAGGG AAGCTCAGGG GGGTTACTGG 1140  
 CTTCTCTTGA GTCACACTGC TAGCAAATGG CAGAACCAAA GCTCAAATAA AAATAAAATA 1200  
 ATTTTCATTC ATTCACCTC

Seq ID NO: 136 Protein sequence:  
 Protein Accession #: NP\_000068.1

1 11 21 31 41 51  
 | | | | |  
 MEPAAGSSME PSADWLATAA ARGRVEEVRA LLEAGALPNA PNSYGRRIPIQ VMMGGSARVA 60  
 ELLLLHGAEP NCADPATLTR PVHDAAREGF LDTLVVLHRA GARLDVRDAW GRLPVDLAEE 120  
 LGHRDVARYL RAAAGGTRGS NHARIDAAEG PSDIPD

Seq ID NO: 137 DNA sequence  
 Nucleic Acid Accession #: NM\_058196.1  
 Coding sequence: 104-421

1 11 21 31 41 51  
 | | | | |  
 TGTGTGGGGG TCTGCTTGGC GGTGAGGGGG CTCTACACAA GCTTCCTTTC CGTCATGCCG 60  
 GCCCCACCC TGGCTCTGAC CATTCTGTTT TCTCTGGCAG GTCATGATGA TGGGCAGCGC 120  
 CCGAGTGGCG GAGCTGCTGC TGCTCCACGG CGCGGAGCCC AACTGCGCCG ACCCCGCCAC 180  
 TCTCACCCGA CCCGTGCACG ACCTGCCCCG GGAGGGCTTC CTGGACACGC TGGTGGTGCT 240  
 GCACCGGGCC GGGGCGCGGC TGGACGTGCG CGATGCCTGG GGCCGTCTGC CCGTGGACCT 300  
 GGCTGAGGAG CTGGGCCATC GCGATGTGCG ACCTGACCTG CGCGCGGCTG CGGGGGGCAC 360  
 CAGAGGCAGT AACCATGCCC GCATAGATGC GCGGAAGGT CCCTCAGACA TCCCGATTG 420  
 AAAGAACCAG AGAGGCTCTG AGAAACCTCG GGAAACTTAG ATCATCAGTC ACCGAAGGTC 480  
 CTACAGGGCC ACAACTGCCC CCGCCACAAC CCACCCCGCT TCTGTAGTTT TCATTTAGAA 540  
 AATAGAGCTT TAAAAATGT CCTGCCTTTT AACGTAGATA TAAGCCTTCC CCCACTACCG 600  
 TAAATGTCCA TTTATATCAT TTTTATATA TTCTTATAAA AATGTAAAAA AGAAAAACAC 660  
 CGCTTCTGCC TTTTCACTGT GTTGGAGTTT TCTGGAGTGA GCACTCACGC CCTAAGCGCA 720  
 CATTTCATGT GGCATTTCTT GCGAGCCTCG CAGCCTCCGG AAGCTGTCGA CTTCATGACA 780  
 AGCATTTTGT GAAGTAGGGA AGTTCAGGGG GGTACTGGC TTCTCTTGAG TCACACTGCT 840  
 AGCAATGGC AGAACCAGAG CTCAAATAAA AATAAAATAA TTTTCATTCA TTCACTC

Seq ID NO: 138 Protein sequence:  
 Protein Accession #: NP\_478103.1

1 11 21 31 41 51  
 | | | | |  
 MMMGSARVAE LLLLHGAEPN CADPATLTRP VHDAAREGFL DTLVVLHRAH ARLDVRDAWG 60  
 RLPVDLAEEEL GHRDVARYLR AAAGGTRGSN HARIDAAEGP SDIPD

Seq ID NO: 139 DNA sequence  
 Nucleic Acid Accession #: NM\_058197.1  
 Coding sequence: 272-684

1 11 21 31 41 51  
 | | | | |  
 CCCAACCTGG GCGCACTTCA GGTGTGCCAC ATTCGCTAAG TGCTCGGAGT TAATAGCACC 60  
 TCCTCCGAGC ACTCGCTCAC GCGGTCCCCT TGCCCTGGAAA GATACCGCGG TCCCTCCAGA 120  
 GGATTTGAGG GACAGGGTCG GAGGGGGCTC TTCCGCCAGC ACCGGAGGAA GAAAGAGGAG 180  
 GGGCTGGCTG GTCACCAGAG GGTGGGGCGG ACCGCGTGCG CTCGGCGGCT GCGGAGAGGG 240  
 GGAGAGCAGG CAGCGGGCGG CGGGGAGCAG CATGGAGCCG GCGGCGGGGA GCAGCATGGA 300  
 GCCGGCGGCG GGGAGCAGCA TGGAGCCTTC GGCTGACTGG CTGGCCACGG CCGCGGCCCCG 360  
 GGGTCGGGTA GAGGAGGTGC GGGCGCTGCT GGAGGCGGGG GCGCTGCCCA ACGCACCGAA 420  
 TAGTTACGGT CGGAGGCCGA TCCAGGTGGT TAGAAGTCT GCAGCGGGAG CAGGGGATGG 480  
 CGGGCGACTC TGGAGGACGA AGTTTGAGG GGAATTGGAA TCAGGTAGCG CTTTCGATTCT 540  
 CCGGAAAAAG GGGAGGCTTC CTGGGGAGTT TTCAGAAGGG GTTTGTAATC ACAGACCTCC 600  
 TCCTGGCGAC GCCCTGGGGG CTTGGGAAAC CAAGGAAGAG GAATGAGGAG CCACGCGCGT 660  
 ACAGATCTCT CGAATGCTGA GAAGATCTGA AGGGGGGAAC ATATTTGTAT TAGATGGAAG 720  
 TCATGATGAT GGGCAGCGCC CGAGTGGCGG AGCTGCTGCT GCTCCACGGC GCGGAGCCCA 780  
 ACTGCGCCCA CCCCGCCACT CTCACCCGAG CCGTGCACGA CGCTGCCCGG GAGGGCTTCC 840  
 TGGACACGCT GGTGGTGTCT CACCGGGCCG GGGCGCGGCT GGACGTGCGC GATGCCTGGG 900  
 GCCGTCTGCC CGTGGACCTG GCTGAGGAGC TGGGCCATCG CGATGTGCGA CCGTACCTGC 960  
 GCGCGGCTGC GGGGGGCACC AGAGGCAGTA ACCATGCCCG CATAGATGCC GCGGAAGGTC 1020  
 CCTCAGACAT CCCCGATTGA AAGAACCAGA GAGGCTCTGA GAAACCTCGG GAACTTAGAT 1080  
 CATCAGTCAC CGAAGGTCCT ACAGGGCCAC AACTGCCCCC GCCACAACCC ACCCCGCTTT 1140  
 CGTAGTTTTC ATTTAGAAA TAGAGCTTTT AAAAATGTCC TGCTTTTAA CGTAGATATA 1200  
 TGCCTTTCCC CACTACCGTA AATGTCCATT TATATCATTT TTTATATATT CTTATAAAAA 1260  
 TGTAATAAAG AAAAAACCCG CTTCTGCCTT TTTACTGTGT TGGAGTTTTC TGGAGTGAGC 1320  
 ACTCACGCCC TAAGCGCACA TTCATGTGGG CATTCTTTCG GAGCCTCGCA GCCTCCGGAA 1380  
 GCTGTCGACT TCATGACAAG CATTTTGTGA ACTAGGGAAG CTCAGGGGGG TTTACTGGCTT 1440  
 CTCTTGAGTC AACTGCTAG CAAATGGCAG AACCAGAGCT CAAATAAAAA TAAATAAATT 1500



Seq ID NO: 140 Protein sequence:  
Protein Accession #: NP\_478104.1

1	11	21	31	41	51	
MEPAAGSSME	PAAGSSMEPS	ADWLATAAAR	GRVEEVRLAL	EAGALPNAPN	SYGRRPIQVG	60
RRSAAGAGDG	GRLWRTKFAG	ELESGSASIL	RKKGRLPGEF	SEGVCNHRPP	PGDALGAWET	120
KEEE						

Seq ID NO: 141 DNA sequence  
Nucleic Acid Accession #: NM\_058195.1  
Coding sequence: 163-684

1	11	21	31	41	51	
CCTCCCTACG	GGCGCCTCCG	GCAGCCCTTC	CCGCGTGCGC	AGGGCTCAGA	GCCGTTCCGA	60
GATCTTGGAG	GTCCGGGTGG	GAGTGGGGGT	GGGGTGGGGG	TGGGGGTGAA	GGTGGGGGGC	120
GGGCGCGCTC	AGGGAAGCGC	GAGTGGCGCC	TGCGGGGCGG	AGATGGGCAG	GGGGCGGTGC	180
GTGGGTCCCA	GTCTGCAGTT	AAGGGGGCAG	GAGTGGCGCT	GCTCACCTCT	GGTGCCAAAG	240
GGCGGCGCAG	CGGCTGCCGA	GCTCGGCCCT	GGAGGCGGCG	AGAACATGGT	GCGCAGGTTT	300
TTGGTGACCC	TCCGGATTCG	GCGCGCGTGC	GGCCCGCCGC	GAGTGAGGGT	TTTCGTGGTT	360
CACATCCCGC	GGCTCACGGG	GGAGTGGGCA	GCGCCAGGGG	CGCCCGCCGC	TGTGGCCCTC	420
GTGCTGATGC	TACTGAGGAG	CCAGCGTCTA	GGGCAGCAGC	CGCTTCCTAG	AAGACCAGGT	480
CATGATGATG	GGCAGCGCCC	GAGTGGCGGA	GCTGCTGCTG	CTCCACGGCG	CGGAGCCCAA	540
CTGCGCCGAC	CCCGCCACTC	TCACCCGACC	CGTGCACGAC	GCTGCCCGGG	AGGGCTTCCT	600
GGACACGCTG	GTGGTGCTGC	ACCGGGCCGG	GGCGCGGCTG	GACGTGCGCG	ATGCCTGGGG	660
CCGTCTGCCC	GTGGACCTGG	CTGAGGAGCT	GGGCCATCGC	GATGTGCGAC	GGTACCTGCG	720
CGCGGCTGCG	GGGGGCACCA	GAGGCAGTAA	CCATGCCCGC	ATAGATGCCG	CGGAAGGTCC	780
CTCAGACATC	CCCGATTGAA	AGAACCAGAG	AGGCTCTGAG	AAACCTCGGG	AAACTTAGAT	840
CATCAGTCAC	CGAAGGTCCT	ACAGGGCCAC	AAC TGCCCCC	GCCACAACCC	ACCCCGCTTT	900
CGTAGTTTTT	ATTTAGAAAA	TAGAGCTTTT	AAAAATGTCC	TGCCTTTTAA	CGTAGATATA	960
TGCCTTCCCC	CACTACCGTA	AATGTCCATT	TATATCATTT	TTTATATATT	CTTATAAAAA	1020
TGTAATAAAG	AAAAACACCG	CTTCTGCCTT	TTCACTGTGT	TGGAGTTTTC	TGGAGTGAGC	1080
ACTCACGCCC	TAAGCGCACA	TTCATGTGGG	CATTTCTTGC	GAGCCTCGCA	GCCTCCGGAA	1140
GCTGTGCACT	TCATGACAA	CATTTTGTGA	ACTAGGGAAG	CTCAGGGGGG	TTACTGGCTT	1200
CTCTTGAGTC	ACACTGCTAG	CAAATGGCAG	AACCAAGGCT	CAAATAAAAA	TAAAATAATT	1260
TTCATTTCATT	CACTC					

Seq ID NO: 142 Protein sequence:  
Protein Accession #: NP\_478102.1

1	11	21	31	41	51	
MGRGRVCVPS	LQLRGQEWRC	SPLVPKGGAA	AAELGPGGGE	NMVRRLVTL	RIRRACGPPR	60
VRVFWVHIPR	LTGEWAAPGA	PAAVALVLM	LRSQRLGQQP	LPRRPGHDDG	QRPSGGAAAA	120
PRRGAQLRRP	RHSHPTARR	CPGGLPGHAG	GAAPGRGAAG	RARCLGPSAR	GPG	

Seq ID NO: 143 DNA sequence  
Nucleic Acid Accession #: NM\_018131  
Coding sequence: 412..1107

1	11	21	31	41	51	
GAAATTGCAC	ACTTAAAGAC	ATCAGTGGAT	GAAATCACAA	GTGGGAAAGG	AAAGCTGACT	60
GATAAAGAGA	GACAGAGACT	TTTGAGAGAA	ATTCGAGTCC	TTGAGGCTGA	GAAGGAGAAG	120
AATGCTTATC	AACTCACAGA	GAAGGACAAA	GAAATACAGC	GACTGAGAGA	CCAAGTGAAG	180
GCCAGATATA	GTACTACCGC	ATTGCTTGAA	CAGCTGGAAG	AGACAACGAG	AGAAGGAGAA	240
AGGAGGGAGC	AGGTGTTGAA	AGCCTTATCT	GAAGAGAAAG	ACGTATTGAA	ACAACAGTTG	300
TCTGCTGCAA	CCTCACGAAT	TGCTGAACCT	GAAAGCAAAA	CCAATACACT	CCGTTTATCA	360
CAGACTGTGG	CTCCAAACTG	CTTCAACTCA	TCAATAAATA	ATATTTCATG	AATGGAAATA	420
CAGCTGAAAG	ATGCTCTGGA	GAAAAATCAG	CAGTGGCTCG	TGTATGATCA	GCAGCGGGAA	480
GTCTATGTAA	AAGGACTTTT	AGCAAAGATC	TTTGAGTTGG	AAAAGAAAAC	GGAAACAGCT	540
GCTCATTAC	TCCACAGCA	GACAAAAAAG	CCTGAATCAG	AAGGTTATCT	TCAAGAAGAG	600
AAGCAGAAAT	GTTACAACGA	TCTCTTGGCA	AGTGCAAAAA	AAGATCTTGA	GGTTGAACGA	660
CAAAACATAA	CTCAGCTGAG	TTTTGAACTG	AGTGAATTTT	GAAGAAAATA	TGAAGAAACC	720
CAAAAAGAAG	TTTCAAAATT	AAATCAGCTG	TTGTATTAC	AAAGAAGGGC	AGATGTGCAA	780
CATCTGGAAG	ATGATAGGCA	TAAAACAGAG	AAGATACAAA	AACTCAGGGA	AGAGAATGAT	840
ATTGCTAGGG	GAAAACTTGA	AGAAGAGAAG	AAGAGATCCG	AAGAGCTCTT	ATCTCAGGTC	900
CAGTCTCTTT	ACACATCTCT	GCTAAAGCAG	CAAGAAGAAC	AAACAAGGGT	AGCTCTGTTG	960
GAACAACAGA	TGCAGGCATG	TACTTTAGAC	TTTGAAAATG	AAAAACTCGA	CCGTCAACAT	1020
GTGCAGCATC	AATTGCATGT	AATCTCTAAG	TGCTCCGAA	AAGCAAGAAA	AAATAACACA	1080
GTTGGAATCC	TTGAAACAGC	TTTCATGAGT	TGCCATCACA	GAGCCATTAG	TCACTTTCCA	1140
AGGAGAGACT	GAAAACAGAG	AAAAAGTTGC	CGCCTCACCA	AAAAGTCCCA	CTGCTGCACT	1200
CAATGGAAGC	CTGGTGGAAT	GTCCCAAGTG	CAATATACAG	TATCCAGCCA	CTGAGCATCG	1260
CGATCTGCTT	GTCCATGTGG	AATACTGTTC	AAAGTAGCAA	AATAAGTATT	TGTTTTGATA	1320
TTAAAAGATT	CAATACTGTA	TTTTCTGTTA	GCTTGTGGGC	ATTTTGAATT	ATATATTTCA	1380
CATTTTGCAT	AAAAGTGCCT	ATCTACCTTT	GACACTCCAG	CATGCTAGTG	AATCATGTAT	1440
CTTTTAGGCT	GCTGTGCATT	TCTCTTGAGC	GTGATACCTC	CCTGACATGG	TTCATCATCA	1500
GGCTGCAATG	ACGAAATGTG	GTCGACGCG	TCTACTGAGA	TACTAACATT	TTGCACTGTC	1560
AAAATACTTG	GTGAGGAAAA	GATAGCTCAG	GTTATTGCTA	ATGGGTAAAT	GCACCAGCAA	1620
GCAAAATATT	TTATGTTTCG	GGGGTTTGA	AAAATCAAAG	ATAATTAAAC	AAGGATCTTA	1680
ACTGTGTTTC	CATTTTTTAT	CCAAGCACTT	AGAAAACCTA	CAATCCTAAT	TTTGATGTCC	1740
ATTGTTAAGA	GGTGGTGATA	GATACTATTT	TTTTTTCATA	TTGTATAGCG	GTTATTAGAA	1800



5 AAGTTGGGGA TTTTCTTGAT CTTTATTGCT GCTTACCATT GAAACTTAAC CCAGCTGTGT 1860  
TCCCCAACTC TGTTCCTGCGC ACGAAACAGT ATCTGTTTGA GGCATAATCT TAAGTGGCCA 1920  
CACACAATGT TTTCTCTTAT GTTATCTGGC AGTAACTGTA ACTGAATTA CATTAGCACA 1980  
TTCTGCTTAG CTAAAATTGT TAAAATAAAC TTTAATAAAC CCATGTAGCC CTCTCATTTG 2040  
ATTGACAGTA TTTTAGTTAT TTTTGGCATT CTAAAGCTG GGCAATGTAA TGATCAGATC 2100  
TTTGTGTTGTC TGAACAGGTA TTTTATACA TGCTTTTGT AAACCAAAAA CTTTAAATT 2160  
TCTTCAGGTT TTCTAACATG CTTACCACTG GGCTACTGTA AATGAGAAAA GAATAAAATT 2220  
ATTTAATGTT TT

10 Seq ID NO: 144 Protein sequence:  
Protein Accession #: NP\_060601

15 1 11 21 31 41 51  
| | | | | |  
MEIQLKDALE KNQQLVLYDQ QREVVYKGLL AKIFELEKKT ETAAHSLPQQ TKKPESEGYL 60  
QEEKQKCYND LLASAKKDL E VERQTITQLS FELSEFRRKY EETQKEVHNL NQLLYSQRRA 120  
DVQHLEDDRH KTEKIQKLE ENDIARGKLE EEKKRSEELL SQVQSLYTS LKQQEEQTRV 180  
20 ALLEQQMQAC TLD FENEKLD RQHVQHQLHV ILKELRKARK NNTVGILETA S

25 Seq ID NO: 145 DNA sequence  
Nucleic Acid Accession #: NM\_001168  
Coding sequence: 50..478

30 1 11 21 31 41 51  
| | | | | |  
CCGCCAGATT TGAATCGCGG GACCCGTTGG CAGAGGTGGC GGCGGCGGCA TGGGTGCCCC 60  
GACGTTGCCC CCTGCGTGGC AGCCCTTTCT CAAGGACCAC CGCATCTCTA CATTCAAGAA 120  
CTGGCCCTTC TTGGAGGGCT GCGCCTGCAC CCCGGAGCGG ATGGCCGAGG CTGGCTTCAT 180  
CCACTGCCCC ACTGAGAACG AGCCAGACTT GGGCCAGTGT TTCTTCTGCT TCAAGGAGCT 240  
GGAAGGCTGG GAGCCAGATG ACGACCCCAT AGAGGAACAT AAAAAGCATT CGTCCGGTTG 300  
CGCTTTCCTT TCTGTCAAGA AGCAGTTTGA AGAATTAACC CTTGGTGAAT TTTTGAAACT 360  
35 GGACAGAGAA AGAGCCAAGA ACAAATTG C AAAGGAAACC AACAATAAGA AGAAAGAATT 420  
TGAGGAAACT GCGAAGAAAG TGCGCCGTGC CATCGAGCAG CTGGCTGCCA TGGATTGAGG 480  
CCTCTGGCCG GAGCTGCCTG GTCCAGAGT GGCTGCACCA CTTCCAGGGT TTATTCCCTG 540  
GTGCCACCAG CCTTCCGTGT GGGCCCTTAG CAATGTCTTA GGAAAGGAGA TCAACATTTT 600  
40 CAAATTAGAT GTTTCAACTG TGCTCCTGTT TTGTCTTGAA AGTGGCACCA GAGGTGCTTC 660  
TGCCTGTGCA GCGGGTGCTG CTGGTAACAG TGGCTGCTTC TCTCTCTCTC TCTCTTTTTT 720  
GGGGGCTCAT TTTTGCTGTT TTGATTCCTG GGGCTTACCAG GTGAGAAGTG AGGGAGGAAG 780  
AAGGCAGTGT CCCTTTTGCT AGAGCTGACA GCTTTGTTCTG CGTGGGCAGA GCCTTCCACA 840  
GTGAATGTGT CTGGACCTCA TGTGTTGAG GCTGTACAG TCCTGAGTGT GGACTTGGCA 900  
45 GGTGCCTGTT GAATCTGAGC TGCAGGTTCC TTATCTGTCA CACCTGTGCC TCCTCAGAGG 960  
ACAGTTTTTT TGTGTTGTG TTTTCTGTT TTTTCTTTT GGTAGATGCA TGACTTGTGT 1020  
GTGATGAGAG AATGGAGACA GAGTCCCTGG CTCCTCTACT GTTAAACAAC ATGGCTTTCT 1080  
TATTTTGT TTGAATGTTAA TTCACAGAAT AGCACAACT ACAATTAATA CTAAGCACAA 1140  
AGCCATTCTA AGTCATTGGG GAAACGGGGT GAACTTCAGG TGGATGAGGA GACAGAATAG 1200  
50 AGTGATAGGA AGCGTCTGGC AGATACTCCT TTTGCCACTG CTGTGTGATT AGACAGGCC 1260  
AGTGAGCCGC GGGGCACATG CTGGCCGCTC CTCCCTCAGA AAAAGGCAGT GGCCTAAATC 1320  
CTTTTAAAT GACTTGGCTC GATGCTGTGG GGGACTGGCT GGGCTGCTGC AGGCCGTGTG 1380  
TCTGTACGCC CAACCTTCAC ATCTGTACAG TTCTCCACAC GGGGGAGAGA CGCAGTCCGC 1440  
CCAGGTCCCC GCTTCTTTG GAGGCAGCAG CTCCCGCAGG GCTGAAGTCT GGCCTAAGAT 1500  
55 GATGGATTTG ATTCGCCCTC CTCCCTGTCA TAGAGCTGCA GGGTGGATTG TTACAGCTTC 1560  
GCTGGAAACC TCTGGAGGTC ATCTCGGCTG TTCCTGAGAA ATAAAAAGCC TGTCATTT

60 Seq ID NO: 146 Protein sequence:  
Protein Accession #: NP\_001159

65 1 11 21 31 41 51  
| | | | | |  
MGAPTLPPAW QPFLKDHRIS TFKNWPFLG CACTPERMAE AGFIHCPTEN EPDLAQCFEC 60  
FKELEGWEPD DDPIEEHKKH SSGCAFLSVK KQFEELTLGE FLKLDREKAK NKIAKETNNK 120  
KKEFEETAKK VRRALQELAA MD

70 Seq ID NO: 147 DNA sequence  
Nucleic Acid Accession #: NM\_014176.1  
Coding sequence: 127-720

75 1 11 21 31 41 51  
| | | | | |  
GCGCGCAGCG CTGGTACCCC GTTGGTCCGC GCGTTGCTGC GTTGTGAGGG GTGTCAGCTC 60  
AGTGCATCCC AGGCAGCTCT TAGTGTGGAG CAGTGAAC TGTTGTGGTT CTCTACTT 120  
GGGATCATGC AGAGAGCTTC ACGTCTGAA AGAGAGCTGC ACATGTTAGC CACAGAGCCA 180  
80 CCCCCAGGCA TCACATGTTG GCAAGATAAA GACCAATGG ATGACCTGCG AGCTCAAATA 240  
TTAGGTGGAG CCAACACACC TTATGAGAAA GGTGTTTTTA AGCTAGAAGT TATCATTCCT 300  
GAGAGGTACC CATTTGAACC TCCTCAGATC CGATTTCTCA CTCCAATTTA TCATCCAAAC 360  
ATTGATTCTG CTGGAAGGAT TTGTCTGGAT GTTCTCAAAT TGCCACCAAA AGGTGCTTGG 420  
85 AGACCATCCC TCAACATCGC AACTGTGTTG ACCTCTATTC AGCTGCTCAT GTCAGAACCC 480  
AACCCTGATG ACCCGCTCAT GGCTGACATA TCCTCAGAAT TTAAATATAA TAAGCCAGCC 540  
TTCTCAAGA ATGCCAGACA GTGGACAGAG AAGCATGCAA GACAGAAACA AAAGGCTGAT 600  
GAGGAAGAGA TGCTTGATAA TCTACCAGAG GCTGGTGACT CCAGAGTACA CAACTCAACA 660  
CAGAAAAGGA AGGCCAGTCA GCTAGTAGGC ATAGAAAAGA AATTTATCC TGATGTTTAG 720  
GGGACTTGTC CTGGTTCATC TTAGTTAATG TGTTCTTTGC CAAGGTGATC TAAGTTGCCT 780  
ACCTTGAATT TTTTTTAAA TATATTTGAT GACATAATTT TTGTGTAGTT TATTTATCTT 840  
GTACATATGT ATTTTGAAAT CTTTAAACC TGAAAAATAA ATAGTCATTT AATGTTGAAA 900



Seq ID NO: 148 Protein sequence:  
 Protein Accession #: NP\_054895.1

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65  
 70

1	11	21	31	41	51	
MQRASRLKRE	LHMLATEPPP	GITCWQDKDQ	MDDLRAQILG	GANTPYEKGV	FKLEVIIPER	60
YFPEPPQIRF	LTPYHPNID	SAGRICLDVL	KLPPKGAWRP	SLNIATVLTS	IQLLMSEPNP	120
DDPLMADISS	EFKYNKPAFL	KNARQWTEKH	ARQKQKADEE	EMLDNLPEAG	DSRVHNSTQK	180
RKASQLVGIE	KKFHPDV					

Seq ID NO: 149 DNA sequence  
 Nucleic Acid Accession #: NM\_003812  
 Coding sequence: 224-2722

20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65  
 70

1	11	21	31	41	51	
TCTCTGCGT	CCCCCCCCCG	GAGTGGCTGC	GAGGCTAGGC	GAGCCGGGAA	AGGGGGCGCC	60
CCCCAGCCCC	GAGCCCCGCG	CCCCGTGCCC	CGAGCCCCGA	CCCCCTGCC	CGCGGCGGCA	120
CCATGCGCGC	CGAGCCGGCG	TGACCGGCTC	CGCCCGCGGC	CGCCCCGCAG	CTAGCCCGGC	180
GCTCTCGCCG	GCCACACGGA	GCGGCGCCCC	GGAGCTATGA	GCCATGAAGC	CGCCCGGCAG	240
CAGCTCGCGG	CAGCCGCCCC	TGGCGGGCTG	CAGCCTTGCC	GGCGCTTCCT	GCGGCCCCCA	300
ACGCGGCCCC	GCCGGCTCGG	TGCCTGCCAG	CGCCCCGGCC	CGCACGCCGC	CCTGCCGCCT	360
GCTTCTCGTC	CTTCTCCTGC	TGCCTCCGCT	CGCCGCCTCG	TCCCGGCCCC	GCGCCTGGGG	420
GGCTGCTGCG	CCCAGCGCTC	CGCATTGGAA	TGAAACTGCA	GAAGGAAATT	TGGGAGTCCT	480
GGCAGATGAA	GACAATACAT	TGCAACAGAA	TAGCAGCAGT	AATATCAGTT	ACAGCAATGC	540
AATGCAGAAA	GAAATCACAC	TGCCTTCAAG	ACTCATATAT	TACATCAACC	AAGACTCGGA	600
AAGCCCTTAT	CACGTTCTTG	ACACAAAGGC	AAGACACCAG	CAAAACATA	ATAAGGCTGT	660
CCATCTGGCC	CAGGCAAGCT	TCCAGATTGA	AGCCTTCGGC	TCCAAATTCA	TTCTTGACCT	720
CATACTGAAC	AATGGTTTGT	TGTCTTCTGA	TTATGTGGAG	ATTCACTACG	AAAATGGGAA	780
ACCACAGTAC	TCTAAGGGTG	GAGAGCACTG	TTACTACCAT	GGAAGCATCA	GAGGCGTCAA	840
AGACTCCAAG	GTGGCTCTGT	CAACCTGCAA	TGGACTTCAT	GGCATGTTTG	AAGATGATAC	900
CTTCGTGTAT	ATGATAGAGC	CACTAGAGCT	GGTTCATGAT	GAGAAAAGCA	CAGGTCGACC	960
ACATATAATC	CAGAAAACCT	TGGCAGGACA	GTATTCTAAG	CAAATGAAGA	ATCTCACTAT	1020
GGAAAGAGGT	GACCAGTGGC	CCTTTCTCTC	TGAATTACAG	TGGTTGAAAA	GAAGGAAGAG	1080
AGCAGTGAAT	CCATCACGTG	GTATATTTGA	AGAAATGAAA	TATTTGGAAC	TTATGATTGT	1140
TAATGATCAC	AAAACGTATA	AGAAGCATCG	CTCTTCTCAT	GCACATACCA	ACAACCTTGC	1200
AAAGTCCGTG	GTCAACCTTG	TGGATTCTAT	TTCAAGGAG	CAGCTCAACA	CCAGGGTTGT	1260
CCTGGTGGCT	GTAGAGACCT	GGACTGAGAA	GGATCAGATT	GACATCACCA	CCAACCCTGT	1320
GCAGATGCTC	CATGAGTTCT	CAAAATACCG	GCAGCGCATT	AAGCAGCATG	CTGATGCTGT	1380
GCACCTCATC	TCGCGGGTGA	CATTTCACTA	TAAGAGAAGC	AGTCTGAGTT	ACTTTGGAGG	1440
TGTCTGTTCT	CGCACAGAG	GAGTTGGTGT	GAATGAGTAT	GGTCTTCCAA	TGGCAGTGGC	1500
ACAAGTATTA	TCGCAGAGCC	TGGCTCAAAA	CCTTGGAATC	CAATGGGAAC	CTTCTAGCAG	1560
AAAGCCAAAA	TGTGACTGCA	CAGAACTCTG	GGGTGGCTGC	ATCATGGAGG	AAACAGGGGT	1620
GTCCCATTCT	CGAAAATTTT	CAAAGTGAGC	CATTTTGGAG	TATAGAGACT	TTTTACAGAG	1680
AGGAGGTGGA	GCCTGCCTTT	TCAACAGGCC	AACAAAGCTA	TTTGAGCCCA	CGGAATGTGG	1740
AAATGGATAC	GTGGAAGCTG	GGGAGGAGTG	TGATTGTGGT	TTTCATGTGG	AATGCTATGG	1800
ATTATGCTGT	AAGAAATGTT	CCCTCTCCAA	CGGGGCTCAC	TGCAGCGACG	GGCCCTGCTG	1860
TAACAATACC	TCATGTCTTT	TTCAGCCACG	AGGGTATGAA	TGCCGGGATG	CTGTGAACGA	1920
GTGTGATATT	ACTGAAATAT	GTACTGGAGA	CTCTGGTCAG	TGCCCAACCA	ATCTTCATAA	1980
GCAAGACGGA	TATGCATGCA	ATCAAAATCA	GGGCCGCTGC	TACAAATGGCG	AGTGCAAGAC	2040
CAGAGACAAC	CAGTGTGAGT	ACATCTGGGG	AACAAAGGCT	GCAGGGTCTG	ACAAGTTCTG	2100
CTATGAAAAG	CTGAATACAG	AAGGCACTGA	GAAGGGAAAC	TGCGGGAAGG	ATGGAGACCG	2160
GTGGATTTCAG	TGCAGCAAAC	ATGATGTGTT	CTGTGGATTC	TTACTCTGTA	CCAATCTTAC	2220
TCGAGCTCCA	CGTATTGGTC	AACTTCAGGG	TGAGATCATT	CCAACTTCCT	TCTACCATCA	2280
AGGCCGGGTG	ATTGACTGCA	GTGGTGCCCA	TGTGTTTATA	GATGATGATA	CGGATGTGGG	2340
CTATGTAGAA	GATGGAAACG	CATGTGGCCC	GTCTATGATG	TGTTTAGATC	GGAAGTGCCT	2400
ACAAATTCAA	GCCCTAAATA	TGAGCAGCTG	TCCACTCGAT	TCCAAGGGTA	AAGTCTGTTC	2460
GGGCCATGGG	GTGTGTAGTA	ATGAAGCCAC	CTGCATTTGT	GATTTACCTT	GGGCAGGGAC	2520
AGATTGCAGT	ATCCGGGATC	CAGTTAGGAA	CCTTCACCCC	CCCAAGGATG	AAGGACCCAA	2580
GGGTCTTAGT	GCCACCAATC	TCATAATAGG	CTCCATCGCT	GGTGCCATCC	TGGTAGCAGC	2640
TATTGTCCTT	GGGGGCACAG	GCTGGGGATT	TAAAAATGTC	AAGAAGAGAA	GGTTCGATCC	2700
TACTCAGCAA	GGCCCCATCT	GAATCGGATG	CGTGGGATGG	ACACCGCCTT	GCACTGTTGG	2760
ATTCTGGGTA	TGACATACTC	GCAGCAGTGT	TACTGGAACT	ATTAAGTTTG	TAAACAAAAC	2820
CTTTGGGTGG	TAATGACTAC	GGAGCTAAAG	TTGGGGTGAC	AAGGATGGGG	TAAAAGAAAA	2880
CTGTCTCTTT	TGGAATAAAT	GTCAAAGAAC	ACCTTTCACC	ACCTGTCAGT	AAACGGGGGA	2940
GGGGGCAAAA	GACCATGCTA	TAAAAAGAAC	TGTTCCAGAA	TCTTTTTTTT	TCCCTAATGG	3000
ACGAAGGAAC	AACACACACA	CAAAAATTAA	ATGCAATAAA	GGAATCATTAA	AAAA	

Seq ID NO: 150 Protein sequence:  
 Protein Accession #: NP\_003803

75  
 80  
 85

1	11	21	31	41	51	
MKPPGSSSRQ	PPLAGCSLAG	ASCGPQRGPA	GSPVASAPAR	TPPCRLLLV	LLLPPLAASS	60
RPRAWGAAAP	SAPHWNETAE	KNLGVLADED	NTLQONSSSN	ISYSNAMQKE	ITLPSRLIYY	120
INQDSESPYH	VLDTKARHQQ	KHNKAVHLAQ	ASFQIEAFGS	KFILDILINN	GLLSSDYVEI	180
HYENGKPKQYS	KGGEHCYHYG	SIRGVKDSKV	ALSTCNGLHG	MFEDDTFVYM	IEPLELVHDE	240
KSTGRPHIIQ	KTLAQYYSKQ	MKNLTMERGD	QWPFLLSELQW	LKRRKRAVNP	SRGIFEEMKY	300
LELMIVNDHK	TYKKHRSSHA	HTNNFAKSVV	NLVDSIYKEQ	LNTRVVLVAV	ETWTEKDQID	360
ITTNVPVQMLH	EFKSKYQRIK	QHADAHLIS	RVTFFHYKRSS	LSYFVGVCSSR	TRGVGVNEYG	420
LPMVAQVLS	QSLAQNLGIQ	WEPSSRKPKC	DCTESWGCCI	MEETGVSHSR	KFSKCSILEY	480
RDFLQRGGGA	CLFNRPTKLF	EPTECGNGYV	EAGEECDGCF	HVECYGLCCK	KCSLSNGAHC	540
SDGPCCNNTS	CLFQPRGYEC	RDAVNECDIT	EYCTGDSGQC	PPNLHKQDGY	ACNQNQGRCY	600
NGECKTRDNQ	CQYIWGTAKA	GSDKFCYEKL	NTEGTEKGNC	GKDGDRWIQC	SKHDFVFCGFL	660



WO 02/086443

PCT/US02/12476

LCTNLTRAPR IGQLQGEIIP TSFYHQGRVI DCSGAHVVDL DDTDVGYVED GTPCGPSMMC 720  
LDRKCLQIQALNMSSCLPDS KGKVCSEGHV CSNEATCICD FTWAGTDCSI RDPVRNLHPP 780  
KDEGPKGPSA TNLIIGSIAG AILVAAIVLG GTGWGFKNVK KRRFDPTQQG PI

5 Seq ID NO: 151 DNA sequence  
Nucleic Acid Accession #: NM\_023915  
Coding sequence: 250-1326

10 1 11 21 31 41 51  
| | | | | |  
GGCACGAGGG TTTCGTTTTT ATGCTTTACC AGAAAATCCA CTTCCTGCCC GACCTTAGTT 60  
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120  
GTGAATGGAC AGCCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180  
15 CCCACGCCCTC AATCGTCCCC AAGTGTTTCC TGACACGCAT CTTTGCTTAC AGTGCATCAC 240  
AACTGAAGAA TGGGGTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCACGGC 300  
CAAGAGAGTC ACAATTCAGG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCCTTCAC 360  
AATGAATTTG ACACAATTGT CTTGCCGGTG CTTTATCTCA TTATATTTGT GGCAAGCATC 420  
TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAAC CAGCTTCATA 480  
20 TTCTATCTCA AAAACATAGT GGTTCAGAGC CTCATAATGA CGCTGACATT TCCATTTTCGA 540  
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTTCTCTG CAGATACACT 600  
TCAGTTTTGT TTTATGCAAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCATT 660  
GATCGCTATC TGAAGGTGGT CAAGCCATTT GGGGACTCTC GGATGTACAG CATAACCTTC 720  
ACGAAGGTTT TATCTGTTTG TGTTCGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780  
25 ATCCTGACAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAAA ACTTAAAGT 840  
CCTTTGGGGG TCAAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900  
GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960  
AGGCAATTCA TAAGTCAGTC AAGCCGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT 1020  
GTGGCTGTGT TTTTACCTG CTTTCTACCA TATCACTTGT GCAGAATTCC TTTTACTTTT 1080  
30 AGTCACTTAG ACAGGCTTTT AGATGAATCT GCACAAAAAA TCCTATATTA CTGCAAAGAA 1140  
ATTACACTTT TCTTGTCTGC GTGTAATGTT TGCCTGGATC CAATAATTTA CTTTTTCATG 1200  
TGTAGGTCAT TTTCAAGAA GCTGTTCAAA AAATCAAATA TCAGAACCAG GAGTGAAAGC 1260  
ATCAGATCAC TGCAAAGTGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTACACTGAT 1320  
GTGTAGGCCT TTTATTGTTT GTTGAATCG ATATGTACAA AGTGTAATAA AATGTTTCTT 1380  
35 TTCATTATCC TTAATAAAAA AA

Seq ID NO: 152 Protein sequence:  
Protein Accession #: NP\_076404

40 1 11 21 31 41 51  
| | | | | |  
MGFNLTLAKL PNNELHQQES HNSGNRSDGP GKNTTLHNEF DTIVLPVLYL IIFVASILLN 60  
GLAVWIFFHI RNKTSFIFYL KNIVVADLIM TLTFPFRIVH DAGFGPWYFK FILCRYTSVL 120  
45 FYANMYTSIV FLGLISIDRY LKVVVKPFQDS RMYSLFTTKV LSVCVWVIMA VLSLPNIILT 180  
NQPTEDNIH DCSKLKSPLG VKWHTATYV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQF 240  
ISQSSRRKRKH NQSIRVVAV PFTCFPLPYHL CRIPFTFSHL DRLLDESAQK ILYYCKEITL 300  
FLSACNVCLD PIIFYFMCRS FSRRLFKKSN ITRSESIRS LQSVRRSEVR IYYDYTDV

50 Seq ID NO: 153 DNA sequence  
Nucleic Acid Accession #: D80008.1  
Coding sequence: 149-739

55 1 11 21 31 41 51  
| | | | | |  
GTTCCGCGCC AAAGCGCGGA GCGGAGGCCG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60  
CGAAAGGAGT GAGGCGCCGA GAGCCCAGAT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120  
AAGGCCGCGG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180  
CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCTTC AACGAGGATG GACTCAGACA 240  
60 AGTTCTGGAG GAGATGAAAG CTTTGTATGA ACAAACCAG TCTGATGTGA ATGAAGCAAA 300  
GTCAGGTGGA CGAAGTGATT TGATACCAAC TATCAAATTT CGACACTGTT CTCTGTTAAG 360  
AAATCGACGC TGCACTGTAG CATACCTGTA TGACCGCTTG CTTCCGGATCA GAGCACTCAG 420  
ATGGGAATAT GGTAGCGTCT TGCCAAATGC ATTACGATTT CACATGGCTG CTGAAGAAAT 480  
GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTCAC TGGGAGGAGA 540  
65 TGAAGGTTTG GGCATTACAC AGGATATGAA ACCACAAAAG AGCCTATATA TTGAAGTCCG 600  
GTGTCTAAAA GACTATGGAG AATTGGAAGT TGATGATGGC ACTTCAGTCC TATTAATAAA 660  
AAATAGCCAG CACTTTTTTAC CTCGATGGAA ATGTGAGCAG CTGATCAGAC AAGGAGTCCT 720  
GGAGCACATC CTGTCTATGAC CATGCGCCGA GGCACCTTCA GGCTTCACTC AACTCATGGA 780  
CTCCTCTGTA CTCACCTCTC CCACCACTCC CTTACCTTCC CTCTTTGATT TTAGAAGCTA 840  
TAGACATTGT TTAAGATAAC TAAGAATACT TGGCTAAGAA GTATAATTG CTAACATATA 900  
70 AGGACTTTCT TTTTAAATG TTGTACACTA TTCTTCTTAC TCTTTTTTGG TTTTGGTTTT 960  
GTTTTGTAGA GACTGTCTCA CTATGTTGCC CAAGCTGGTC TCAAACCTCT GGCCTCAAGC 1020  
AGTCTCCCA CCTTAGCTTC TCAAAGTGTG GAGATCACAG GCGTGAGCCA CTGCACCCGG 1080  
CCCCTACTCC TTTTCTAAT AAGCTGTATC TGTAATCACA GCATTCTTAC AGTTGTTACA 1140  
75 GTGTGTTTTT TAAATGAAAG TAAACATGGT TACATTTGAA TCTCTTAAAT AAGCAGTCAC 1200  
TTGGCTGGAC AGGAAGAAGG TAGATCCTGT GTGTCTTGTG TTCTGGTCAT GTGTATTGTA 1260  
CAAGCTAGAG AGCTGAATTT CTGAGATAGA CATTTTCAAA TCACATGCAA GTGAAGATGA 1320  
TGGTCTGTAG AAATTTTCAG TATATATAAT GTTTAATGAC ATACTAATTT ATCATCTGGC 1380  
TATTTGGGAA GGAAGGACAC ACATGGATTG TGCACATTTT CACCATGGTG GCTGGTGTGG 1440  
CTTGTGGCTA TGGGGTGATC ACCAGTATCA CCACTTTGGA AGGGGACAGT GAAATTGGGG 1500  
80 CTAGAGAAGG AACTTTGTAC AGTTTCCCTT GAGATTGAGA TTGACTGAAA AGTCACATGA 1560  
AGAGTTGATT GTCTTTTAAAT GGTATGTTTT AAACAGCTGA CATTTTAAAT TTTGATGAAA 1620  
TCCAGTTTAT TCGTTTGTTC TTTTATGCTT TGGGCTTTGC ATCCGAGAAA TCTTTTCCCA 1680  
TCCCAAGATC ACAATTTTTT TTCTTTTTTA CTTCTAGAAG TGTTATAATT TTAAGCTTTA 1740  
TACTTTGGTC TATGACCCGT TTTTTTTTTT GTTTTGTGTT GTTTTTTCGT TTGTTTCTTT 1800  
85 GTTTTGGAGT GGAGTCTTGT TCTGTACCCC AGGCTGGGGT GCAGTGGCGT GATCTTGGCT 1860  
CACTGCAATC TCTATCCCTT GGGTTCAGGT GATTCTCTTG TCTCAGCCTC CCAAGTAGCT 1920  
GGGATTACAG GCACAGGCCG CCACGCCTGG CTAATTTTTG TATTTTTTAGT AGAGACAGAG 1980



TTTTACCATG TTGGCCAGGC TGGTTTCAAA CTCCTGACCT CAAGTGACCC ACCTTGGCCT 2040  
 CCCAAAGTTT TGGGATTACA AGTGTGGGCC ACCGCGGCCA GCCTATGATC CATTTTGAAT 2100  
 GAATTTTTTA TATGGTGCAA GGTGTCAATC CACCTTCACT TTTTCTTGGG AATATAGATA 2160  
 TCCAGCTGTT TCACTACCAT TTTTGTAAAG GACTGCCCTT TGCTCTATCA CCTTTGCATT 2220  
 TTTGTTAAAA AGTAGTTGTC AATGTATATG TGGGTTTATT TCAGGACTCT GTTTTGTTC 2280  
 ATTGACCTGT TTTTCTCTCC TGAATGCCAA TACCATATTT GTATGTAGTG TATGTAATTT 2340  
 TCTAATAATT CTTGAAACAG ATAGTATTAA TGTGTCATAT TTTTGCTGTT GTTTGTATTT 2400  
 TTTGTAGAGA TGGGGTTTCA CCGTGTGGGC CAGGCTGTGT TGAACCTCTG AGCTAAAGCA 2460  
 ATACACTTGC CTCGTCTCTC CCATGTGCTG GGATTACAGG CGTGAGCCTT GGTGCTGGCC 2520  
 CAGTGTACCA CATTTCTTTT TGAGATTTGT TTTGGCTATG TTAAGTCCTT TGCTTTTGAT 2580  
 GTGAAATTTG GGAACAGGCA GGGTGTGGTG GCTTATGCCT GTAATCCTAG AACTTTGGGA 2640  
 GGCCTAGATG GGTGGATCAC TTGAGCTCAG GAGTTCAGAG CCAGCCCGGG CCTATGGCAA 2700  
 AACTCCGTCT CTACAAAAAA TAGAAAAAAT TAGCCAGGTG TGGTGGTGCA TGCCTGTAGT 2760  
 CACAGTTACA CGCAGGCTG AGGTGGGAGG ATCACTTGAA CCCAGAGGT CAAGACTGCA 2820  
 GTGAGCTGAG ATCACACCAC TGTACTCCAG CCTGGGTGAC AAAGTGAGAC TCTATCTCAA 2880  
 AAAGAAATTA GGATCAATTT GTCAATTTCT ACAACAACAA CAACAAAAAC CCCTGTTGGG 2940  
 CACCTTGATT GAGATTGCAT TGAATTTATA TAAACTGTG GGGAGAATTG ACATCTTAAT 3000  
 AATATTGAGT CTTCTGGCCT ATAAACAAGG TCTGTCTTCC TAGGTATTAA TGTTTTGTCT 3060  
 TCTATTTCTC TTAATAATCT TTTGTAGTTT TCACTGTACA GGTCTACCAT GTCAGCATTT 3120  
 CATAGTTTTG ATGCTAAATG GTATTTTAAA ATTTCAAATT CTAACCACTT GTTGCTAGTA 3180  
 AATAGAAATA CAATTGATGT TGAACCTGTA TCCTTCAGCC TTGCTAAACT GTGAGTTCTC 3240  
 ATGGTGTTTT TGTAATATTAC ATCAACAGTC ATGTGTTCTA TGAATAAAGA GTTTTACTCC 3300  
 TTC

Seq ID NO: 154 Protein sequence:  
Protein Accession #: BAA11503.1

1 11 21 31 41 51  
 MFCEKAMELI RELHRAPEGQ LPAFNEDGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60  
 TIKFRHCSLL RNRRTVAYL YDRLLRIRAL RWEYGSVLPN ALRFHMAAEE MEWFNNYKRS 120  
 LATYMRSLGG DEGLDITQDM KPPKSLYIEV RCLKDYGEFE VDDGTSVLLK KNSQHFLPRW 180  
 KCEQLIRQGV LEHILS

Seq ID NO: 155 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 149-709

1 11 21 31 41 51  
 GTTCGGCGCC AAAGCGCGGA GCGGAGGCCG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60  
 CGAAAGGAGT GAGGCGCCGA GAGCCCAGAT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120  
 AAGGCCGCGG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180  
 CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCCTC AACGAGGATG GACTCAGACA 240  
 AGTTCTGGAG GAGATGAAAG CTTTGTATGA ACAAACCAG TCTGATGTGA ATGAAGCAAA 300  
 GTCAGGTGGA CGAAGTGATT TGATACCAAC TATCAAATTT CGACACTGTT CTCTGTTAAG 360  
 AAATCGACGC TGCACTGTAG CATACCTGTA TGACCGCTTG CTTCCGATCA GAGCACTCAG 420  
 ATGGGAATAT GGTAGCGTCT TGCCAAATGC ATTACGATTT CACATGGCTG CTGAAGAAAT 480  
 GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTCAC TGGGAGGAGA 540  
 TGAAGGTTTG GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGCTGG 600  
 ATGCAGTGGC GCGATCTCGG CTCAACCTGC AACCTCCACC TCCAGGTTT ACCTCAACTG 660  
 CAACCTCCAC CTCCCAGGTC CCGTGTCTAA AAGACTATGG AGAATTTGAA GTTGATGATG 720  
 GCACTTCAGT CCTATTAAAA AAAAATAGCC AGCACTTTTT ACCTCGATGG AAATGTGAGC 780  
 AGCTGATCAG ACAAGGAGTC CTGGAGCACA TCCTGTCTAT ACCATGCGCC GAGGCACTTC 840  
 CAGGCTTCAC TCAACTCATG GACTCCTCTG TACTCACTCT CTCCACCACT CCCTTCACCT 900  
 CCCTCTTTGA TTTTAGAAGC TATAGACATT GTTTAAGATA ACTAAGAATA CTTGGCTAAG 960  
 AAGTATAATT TGCTAACTAT TAAGGACTTT CTTTTTTTAA TGTGTACAC TATTCTTCCT 1020  
 ACTCTTTTTT GGTTTTGGTT TTGTTTTGTA GAGACTGTCT CACTATGTTG CCAAGCTGG 1080  
 TCTCAAACCT CTGGCCTCAA GCAGTCCTCC CACCTTAGCT TCTCAAAGTG TTGAGATCAC 1140  
 AGGCGTGAGC CACTGCACCC GGGCCCTACT CTTTTTCTA ATAAGCTGTA TCTGTAATCA 1200  
 CAGCATTCCT ACAGTTGTTA CAGTGTGTTT TTTAAATGAA AGTAAACATG GTTACATTTG 1260  
 AATCTCTTAA ATAAGCAGTC ACTTGGCTGG ACAGGAAGAA GGTAGATCCT GTGTGTCTTG 1320  
 TTTTCTGGTC ATGTGTATTG TACAAGCTAG AGAGCTGAAT TTCTGAGATA CACATTTTCA 1380  
 AATCACATGC AAGTGAAGAT GATGGTCTGT AGAAATTTTC AGTATATATA ATGTTTAAATG 1440  
 ACATACTAAT TTATCATCTG GCTATTTGGG AAGGAAGGAC ACACATGGAT TTTGCACATT 1500  
 TCCACCATGG TGGCTGGTGT GGCTGTGGC TATGGGGTGA TCACCAGTAT CACCACTTTG 1560  
 GAAGGGGACA GTGAATTTGG GGCTAGAGAA GGAACTTTGT ACAGTTTTC CTGAGATTCA 1620  
 GATTGACTGA AAAGTCACAT GAAGAGTTGA TTGTCTTTTA ATGGTATGTT TTAACAGCT 1680  
 GACATTTTAA ATTTTGATGA AATCCAGTTT ATTCGTTTGT TCTTTTATGC TTTGGGTGTT 1740  
 GCATCCGAGA AATCTTTTCC CATCCCAAGA TACAATTTT TTTTCTTTT TACTTCTAGA 1800  
 AGTGTTATAA TTTTAAAGCT TATACTTTGG TCTATGACCC GTTTTTTTTT TTGTTTTGTT 1860  
 TTGTTTTTTC GTTTGTTTCT TTGTTTTGAG ATGGAGTCTT GTTCTGTAC CCAGGCTGGG 1920  
 GTGCAGTGGC GTGATCTTGG CTCAGTCAA TCTCTATCCC CTGGGTCAA GTGATTCTCT 1980  
 TGTCTCAGCC TCCCAAGTAG CTGGGATTAC AGGCACAGGC CGCCACGCCT GGCTAATTTT 2040  
 TGTAATTTTA GTAGAGACAG AGTTTAACTA TTTGGCCAG GCTGGTTTCA AACTCCTGAC 2100  
 CTGATTTGAC CCACCTGGC CTCCCAAGT TTTGGGATTA CAAGTGTGGG CCACCGCGGC 2160  
 CAGCCTATGA TCCATTTTGA ATGAATTTT TATATGGTGC AAGGTGTCAA TCCACCTTCA 2220  
 CTTTTTCTTG GGAATATAGA TATCCAGCTG TTTCACTACC ATTTTTTGAA AGGACTGCCC 2280  
 TTTGCTCTAT CACCTTTGCA TTTTGTGTTA AAAGTAGTTG TCAATGTATA TGTGGGTTTA 2340  
 TTTCAGGACT CTGTTTGTGTT CCATTGACCT GTTTTTCTCT CCTGAATGCC AATACCATAT 2400  
 TTGTATGTAG TGTATGTAAT TTTCTAATAA TTCTTGAAC AGATAGTATT AATGTGTCTAT 2460  
 ATTTTGTCTG TTGTTGTATG TTTTGTAGA GATGGGTTT CACCGTGTG GCCAGGCTGT 2520  
 GTTGAACCTC TGAGCTAAAG CAATACACTT GCCTCGTCTT CCCCATGTGC TGGGATTACA 2580  
 GCGGTGAGCC TTGGTGTCTG CCCAGTGTAC CACATTTCTT TTTGAGATT GTTTTGGCTA 2640  
 TGTAAAGTCC TTTGCTTTTG ATGTGAAATT TGGGAACAGG CAGGGTGTGG TGGCTTATGC 2700  
 CTGTAATCCT AGAAGTTTGG GAGGCTAGA TGGGTGGATC ACTTGAGCTC AGGAGTTCCA 2760  
 GACCAGCCCG GGCCTATGGC AAAACTCCGT CTCTACAAAA AATAGAAAAA ATTAGCCAGG 2820



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80

TGTGGTGGTG	CATGCCTGTA	GTCACAGTTA	CACGGCAGGC	TGAGGTGGGA	GGATCACTTG	2880
AACCCAGAG	GTCAAGACTG	CAGTGAGCTG	AGATCACACC	ACTGTACTCC	AGCCTGGGTG	2940
ACAAAGTGAG	ACTCTATCTC	AAAAAGAAAT	TAGGATCAAT	TTGTCAATTT	CTACAACAAC	3000
AACAACAAAA	ACCCTGTGTT	GGCACCTTGA	TTGAGATTGC	ATTGAATTTA	TATAAACTG	3060
TTGGGAGAAT	TGACATCTTA	ATAATATTGA	GTCTTCTGGC	CTATAAACAA	GGTCTGTCTT	3120
CCTAGGTATT	AATGTTTGT	CTTCTATTTT	TCTTAATAAT	CTTTTGTAGT	TTTCAGTGTG	3180
CAGGTCTACC	ATGTCAGCAT	TTCATAGTTT	TGATGCTAAA	TGGTATTTTA	AAATTTCAAA	3240
TTCTAACCAC	TTGTTGCTAG	TAAATAGAAA	TACAATTGAT	GTTGAACTTG	TATCCTTCAG	3300
CCTTGCTAAA	CTGTGAGTTC	TCATGGTGTT	TTTGTAATAAT	ACATCAACAG	TCATGTGTTC	3360
TATGAATAAA	GAGTTTACT	CCTTC				

Seq ID NO: 156 Protein sequence:  
Protein Accession #: Eos sequence

15  
20

1	11	21	31	41	51	
MFCEKAMELI	RELHRAPEGQ	LPAFNEDGLR	QVLEEMKALY	EQNQSDVNEA	KSGGRSDLIP	60
TIKFRHCSLL	RNRRTVAYL	YDRLLRIRAL	RWEYGSVLPN	ALRFHMAAEE	MEWFNNYKRS	120
LATYMRSLGG	DEGLDITQDM	KPPKSLYIEA	GCSGAISAQP	ATSTSQVHLN	CNLHLPGPVS	180
KRLWRI						

Seq ID NO: 157 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 148-621

25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80

1	11	21	31	41	51	
TTCGGCGCCA	AAGCGCGGAG	CGGAGGCCGA	GGCGAGAGCC	TGGCGCTGTA	GGACTAGAAC	60
GAAAGGAGTG	AGGCGCCGAG	AGCCCAGATA	CCATTTTGGC	GTGAGAGCTG	GTGGTTGGCA	120
AGGCCGCGGG	AGTGGGAAGC	GTCCGCCATG	TTCTGCGAAA	AAGCCATGGA	ACTGATCCGC	180
GAGCTGCATC	GCGCGCCCGA	AGGGCAACTG	CCTGCCTTCA	ACGAGGATGG	ACTCAGACAA	240
GTTCTGGAGG	AGATGAAAGC	TTTGTATGAA	CAAAACCAGT	CTGATGTGAA	TGAAGCAAAG	300
TCAGGTGGAC	GAAGTGATTT	GATACCAACT	ATCAAATTTT	GACACTGTTC	TCTGTTAAGA	360
AATCGACGCT	GCACTGTAGC	ATACCTGTAT	GACCGCTTGC	TTCGGATCAG	AGCACTCAGA	420
TGGGAATATG	GTAGCGTCTT	GCCAAATGCA	TTACGATTTT	ACATGGCTGC	TGAAGAAGTC	480
CGGTGTCTAA	AAGACTATGG	AGAATTTGAA	GTGATGATG	GCACTTCAGT	CCTATTAATA	540
AAAAATAGCC	AGCACTTTTT	ACCTCGATGG	AAATGTGAGC	AGCTGATCAG	ACAAGGAGTC	600
CTGGAGCACA	TCCTGTCTAT	ACCATGCGCG	GAGGCACCTT	CAGGCTTCAC	TCAACTCATG	660
GACTCCTCTG	TACTCACTCT	CTCCACCACT	CCCTTCACCT	CCCTCTTTGA	TTTTAGAAGC	720
TATAGACATT	GTTTAAGATA	ACTAAGAATA	CTTGGCTAAG	AAGTATAATT	TGCTAACTAT	780
TAAGGACTTT	CTTTTTTTAA	TGTTGTACAC	TATTCTTCCT	ACTCTTTTTT	GGTTTTGGTT	840
TTGTTTTGTA	GAGACTGTCT	CACTATGTTG	CCCAAGCTGG	TCTCAAACCT	CTGGCCTCAA	900
GCAGTCCTCC	CACCTTAGCT	TCTCAAAGTG	TTGAGATCAC	AGGCGTGAGC	CACTGCACCC	960
GGCCCCCTACT	CCTTTTCTTA	ATAAGCTGTA	TCTGTAATCA	CAGCATTCCT	ACAGTTGTTA	1020
CAGTGTGTTT	TTTAAATGAA	AGTAAACATG	TTGTACATTT	AATCTCTTAA	ATAAGCAGTC	1080
ACTTGGCTGG	ACAGGAAGAA	GGTAGATCCT	GTGTGTCTTG	TTTTCTGGTC	ATGTGTATTG	1140
TACAAGCTAG	AGAGCTGAAT	TTCTGAGATA	CACATTTTCA	AATCACATGC	AAGTGAAGAT	1200
GATGGTCTGT	AGAAATTTTC	AGTATATATA	ATGTTTAAAT	ACATACTAAT	TTATCATCTG	1260
GCTATTTGGG	AAGGAAGGAC	ACACATGGAT	TTTGCACATT	TCCACCATGG	TGGCTGGTGT	1320
GGCTTGTGGC	TATGGGGTGA	TCACCAGTAT	CACCAGTTTG	GAAGGGGACA	GTGAAATTGG	1380
GGCTAGAGAA	GGAACTTTGT	ACAGTTTTC	CTGAGATTCA	GATTGACTGA	AAAGTCACAT	1440
GAAGAGTTGA	TTGTCTTTTA	ATGGTATGTT	TTAAACAGCT	GACATTTTAA	ATTTTGATGA	1500
AATCCAGTTT	ATTCGTTTGT	TCTTTTATGC	TTTGGGTGTT	GCATCCGAGA	AATCTTTTCC	1560
CATCCCAAGA	TCACAAATTT	TTTTCTTTT	TACTTCTAGA	AGTGTTATAA	TTTTAAGCTT	1620
TATACTTTGG	TCTATGACCC	GTTTTTTTTT	TTGTTTTGTT	TTGTTTTTTC	GTTTTGTTCT	1680
TTGTTTTGAG	ATGGAGTCTT	GTTCTGTAC	CCAGGCTGGG	GTGCAGTGGC	GTGATCTTGG	1740
CTCCTGCAA	TCTCTATCCC	CTGGGTTCAA	GTGATTCTCT	TGCTCAGCC	TCCCAAGTAG	1800
CTGGGATTAC	AGGCACAGGC	CGCCACGCCT	GGCTAATTTT	TGTATTTTAA	GTAGAGACAG	1860
AGTTTTACCA	TGTTGGCCAG	GCTGGTTTCA	AACTCCTGAC	CTCAAGTGAC	CCACCTTGGC	1920
CTCCCAAAGT	TTTGGGATTA	CAAGTGTGGG	CCACCGCGGC	CAGCCTATGA	TCCATTTTGA	1980
ATGAATTTTT	TATATGTTGC	AAGGTGTCAA	TCCACCTTCA	CTTTTCTTGG	GGAATATAGA	2040
TATCCAGCTG	TTTCACTACC	ATTTTTTGAA	AGGACTGCCC	TTTGCTCTAT	CACCTTTGCA	2100
TTTTTGTTAA	AAAGTAGTTG	TCAATGTATA	TGTGGGTTTA	TTTCAGGACT	CTGTTTTGTT	2160
CCATTGACCT	GTTTTTCTCT	CCTGAATGCC	AATACCATAT	TTGTATGTAG	TGTATGTAAT	2220
TTTCTAATAA	TTCTTGAAAC	AGATAGTATT	AATGTGTCAT	ATTTTTGCTG	TTGTTTGTAT	2280
TTTTTGTAGA	GATGGGGTTT	CACCGTGTGG	GCCAGGCTGT	GTTGAACTCC	TGAGCTAAAG	2340
CAATACACTT	GCCTCGTCCT	CCCCATGTGC	TGGGATTACA	GGCGTGAGCC	TTGGTGCTGG	2400
CCCAGTGATC	CACATTTCTT	TTTGAGATTT	GTTTTGGCTA	TGTTAAGTCC	TTTGCTTTTG	2460
ATGTGAAATT	TGGGAACAGG	CAGGGTGTGG	TGGCTTATGC	CTGTAATCCT	AGAAGTTTGG	2520
GAGGCTAGAA	TGGGTGATC	ACTTGAGCTC	AGGAGTTCCA	GACCAGCCCG	GGCCTATGGC	2580
AAAACCTCCG	CTCTACAAAA	AATAGAAAAA	ATTAGCCAGG	TGTGGTGGTG	CATGCCTGTA	2640
GTCACAGTTA	CACGGCAGGC	TGAGGTGGGA	GGATCACTTG	AACCCAGAG	GTCAAGACTG	2700
CAGTGAGCTG	AGATCACACC	ACTGTACTCC	AGCCTGGGTG	ACAAAGTGAG	ACTCTATCTC	2760
AAAAAGAAAT	TAGGATCAAT	TTGTCAATTT	CTACAACAAC	AACAACAAAA	ACCCCTGTTG	2820
GGCACTTTGA	TTGAGATTGC	ATTGAAATTA	TATAAACTG	TTGGGAGAAT	TGACATCTTA	2880
ATAATATTGA	GTCTTCTGTC	CTATAAACAA	GGTCTGTCTT	CCTAGGTATT	AATGTTTTGT	2940
CTTCTATTTT	TCTTAATAAT	CTTTTGTAGT	TTTCAGTGTG	CAGGTCTACC	ATGTCAGCAT	3000
TTCATAGTTT	TGATGCTAAA	TGGTATTTTA	AAATTTCAAA	TTCTAACCAC	TTGTTGCTAG	3060
TAAATAGAAA	TACAATTGAT	GTTGAACTTG	TATCCTTCAG	CCTTGCTAAA	CTGTGAGTTC	3120
TCATGGTGTT	TTTGTAATAAT	ACATCAACAG	TCATGTGTTC	TATGAATAAA	GAGTTTTACT	3180
CCTTC						

Seq ID NO: 158 Protein sequence:  
Protein Accession #: Eos sequence

85

1	11	21	31	41	51



MFCEKAMELI RELHRAPEGQ LPAFNEDGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60  
TIKFRHCSLL RNRRTVAYL YDRLLRIRAL RWEYGSVLPN ALRFHMAAEE VRCLKDYGEF 120  
EVDDGTSVLL KKNQSHFLPR WKCEQLIRQG VLEHILS

5

Seq ID NO: 159 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 149-229

10

1	11	21	31	41	51	
GTTCGGCGCC	AAAGCGCGGA	GCGGAGGCCG	AGGCGAGAGC	CTGGCGCTGT	AGGACTAGAA	60
CGAAAGGAGT	GAGGCGCCGA	GAGCCCAGAT	ACCATTTTGG	CGTGAGAGCT	GGTGGTTGGC	120
AAGGCCGCGG	GAGTGGGAAG	CGTCCGCCAT	GTTCTGCGAA	AAAGCCATGG	AACTGATCCG	180
CGAGCTGCAT	CGCGCGCCCC	AAGGGCAACT	GCCTGCCTTC	AACAATTAGC	TGGGTGTGGT	240
GGCACACACC	TGTAGTCCCA	GCAACTTAGG	AGGCTGAAGT	GAGAGGATTG	CATGGCTCCA	300
GGAAAGTTGAA	ACTGCAGTGA	ACTGTGGTCA	CGCTATTACA	CTCCAGCCTG	GGTGACAGAC	360
TGAATCCCTG	TCTCAAAAAG	GAAAAGGAGG	ATGGACTCAG	ACAAGTTCTG	GAGGAGATGA	420
AAGCTTTGTA	TGAACAAAAC	CAGTCTGATG	TGTTCTCTGT	TAAGAAATCG	ACGCTGCACT	480
GTAGCATACC	TGTATGACCG	CTTGCTTCGG	ATCAGAGCAC	TCAGATGG		

20

Seq ID NO: 160 Protein sequence:  
Protein Accession #: Eos sequence

25

1	11	21	31	41	51	
ATGTTCTGCG	AAAAAGCCAT	GGAAGTGATC	CGCGAGCTGC	ATCGCGCGCC	CGAAGGGCAA	60
CTGCCTGCCT	TCAACAATTA	G				

30

Seq ID NO: 161 DNA sequence  
Nucleic Acid Accession #: U10694  
Coding sequence: 1333-2280

35

1	11	21	31	41	51	
GGATCCGGCC	GGATCTCAGG	GAGGTGAGGA	CTTTGTTCTC	AGAGGGTGTG	TGTGGACAAA	60
ACAGGGAGGC	CCTGTGTTTC	ACAGACACAG	TGGTCCCAGG	ATTGGAGAGC	AGTCCAGGTG	120
AGGAACCTAA	GGGAGGATCG	AGGGTACCTC	CAGGCCAGAG	AAACTCTCAG	ATCAAGAGAG	180
TTTGCCCTGC	CCCTACTGTC	ACCCACAGAG	GCCCGGGCAG	GGCTGTCTGC	TGAGGTCCCT	240
CCTTTATCCT	GGGATCACTG	GTGTCGGGGA	GGGCTGGCCT	TGGTCTGAGG	GGGCTGCACT	300
CACGTACAGC	GAGGGAGGGT	CCCAGGCCCT	GCCAGGAGTC	CAGGTGCAGA	CTGAGGGGAC	360
CCCCTCACC	AAACACAGAG	GACCTAGCCC	CACCCTGCCC	CTTGTGTCTC	CTGAGGGGAG	420
CCGCTGGGTG	GATGGACTCC	CCTCACTTCC	TCTTCAGGTG	TCTCTGGAG	ATAGGGCCTC	480
AGGTCAACAG	AGGGAGGGTT	CCAGACCCCT	CAGGCATCAA	GATGAGGACC	AGGCAGTATC	540
CTCACCCAG	GACACATGGA	CCCCATTGAA	TTTAGACATC	TCTTACTGTA	CTTCCGAGGA	600
AACCCTGGGC	AGGTGTGGGC	AGATGTTGGT	TGGGGCATGT	CCTTCTGTTC	CATATCAGGG	660
ATGTGAGCTC	CTGATCTGAG	AGACTCTCAG	GCAAGTAGAG	GAGTAGAGTC	CAGTCCCTGC	720
CAGGAGAAAG	GTCAGGGCCC	TGAGTGAGCG	CAGAGGGGAC	CATCCACCCC	AAAAGTGTGT	780
AGAACTCAAG	AGTGTCCAGC	CCGCCCTCTT	GACAGCACTG	AGGGACCGGG	GCTCTGCCTG	840
CAGTCTGCAG	CCTAAGGGCC	CCTCGATTCC	TCTTCCAGGA	GCTCCAGGAA	GCAGGCAGGC	900
CTTGGTCTGA	GACAGTGTCC	TCAGGTGCGA	GAGCAGAGGA	GACCCAGGCA	GTGTCTCAGC	960
TGAAGGTGAA	GTGTTCACCC	TGAATGTGCA	CCAAGGGCCC	CACCTGCCCC	AGCACACATG	1020
GGACCCCAT	GCACCTGGCC	CCATTCCCCC	TACTGTCACT	CATAGAGCCT	TGATCTCTGC	1080
AGGCTAGCTG	CACGCTGAGT	AGCCCTCTCA	CTTCTCCCTT	CAGGTCTCTG	GGACAGGCTA	1140
ACCAGGAGGA	CAGGAGCCCC	AAGAGGCCCC	AGAGCAGCAC	TGACGAAGAC	CTGTAAGTCA	1200
GCCTTTGTTA	GAACCTCCAA	GGTTCGGTTC	TCAGCTGAAG	TCTCTCACAC	ACTCCCTCTC	1260
TCCCCAGGCC	TGTGGGTCTC	CATCGCCCAG	CTCCTGCCCA	CGCTCCTGAC	TGCTGCCCTG	1320
ACCAGAGTCA	TCATGCTCTT	CGAGCAGAGG	AGTCCGCACT	GCAAGCCTGA	TGAAGACCTT	1380
GAAGCCCAAG	GAGAGGACTT	GGGCCTGATG	GGTGACACAG	AACCCACAGG	CGAGGAGGAG	1440
GAGACTACCT	CCTCCTCTGA	CAGCAAGGAG	GAGGAGGTGT	CTGCTGCTGG	GTCTATCAAGT	1500
CCTCCCCAGA	GTCTCTCAGG	AGGCGCTTCC	TCCTCCATT	CCGTCTACTA	CACCTTATGG	1560
AGCCAATTCT	ATGAGGGCTC	CAGCAGTCAA	GAAGAGGAAG	AGCCAAGCTC	CTCGGTCTGAC	1620
CCAGCTCAGC	TGGAGTTCAT	GTTCCAAGAA	GCACTGAAAT	TGAAGGTGGC	TGAGTTGGTT	1680
CATTTCTCTG	TCCACAAATA	TCGAGTCAAG	GAGCCGGTCA	CAAAGGCAGA	AATGCTGGAG	1740
AGCGTCATCA	AAAATTACAA	GCGCTACTTT	CCTGTGATCT	TCGGCAAAGC	CTCCGAGTTC	1800
ATGCAGGTGA	TCTTTGGCAC	TGATGTGAAG	GAGGTGGACC	CCGCCGGCCA	CTCCTACATC	1860
CTTGCTCACTG	CTCTTGGCCT	CTCGTGCGAT	AGCATGCTGG	GTGATGGTCA	TAGCATGCCC	1920
AAGGCCGCC	TCCTGATCAT	TGCTCTGGGT	GTGATCCTAA	CCAAAGACAA	CTGCGCCCTT	1980
GAAGAGGTTA	TCTGGGAAGC	GTTGAGTGTG	ATGGGGGTGT	ATGTTGGGAA	GGAGCACATG	2040
TTCTACGGGG	AGCCCAGGAA	GCTGCTCACC	CAAGATTGGG	TGCAGGAAAA	CTACCTGGAG	2100
TACCGGCAGG	TGCCCCGCGC	TGATCCTGCG	CACTACGAGT	TCCTGTGGGG	TTCCAAGGCC	2160
CACGCTGAAA	CCAGCTATGA	GAAGGTGATA	AATTATTTGG	TCATGCTCAA	TGCAAGAGAG	2220
CCCCTCTGCT	ACCCATCCCT	TTATGAAGAG	GTTTTGGGAG	AGGAGCAAGA	GGGAGTCTGA	2280
GCACCAGCCG	CAGCCGGGGC	CAAAGTTTGT	GGGGTCAGGG	CCCCATCCAG	CAGCTGCCCT	2340
GCCCCATGTG	ACATGAGGCC	CATTCTTCGC	TCTGTGTTTG	AAGAGAGCAA	TCAGTGTTCT	2400
CAGTGGCAGT	GGGTGGAAGT	GAGCACACTG	TATGTCTCTT	CTGGGTTCTT	TGTCTATTGG	2460
GTGATTTGGA	GATTTATCCT	TGCTCCCTTT	TGGAATTGTT	CAAATGTTCT	TTTAATGGTC	2520
AGTTTAATGA	ACTTCACCAT	CGAAGTTAAT	GAATGACAGT	AGTCACACAT	ATTGCTGTTT	2580
ATGTTATTTA	GGAGTAAGAT	TCTTGCTTTT	GAGTCACATG	GGGAAATCCC	TGTTATTTTG	2640
TGAATTGGGA	CAAGATAACA	TAGCAGAGGA	ATTAATAATT	TTTTTGAAAC	TTGAACTTAG	2700
CAGCAAAATA	GAGCTCATAA	AGAAATAGTG	AAATGAAAT	GAGTTAATT	CTGCTTAT	2760
ACCTCTTTCT	CTCTCTGTGA	AAATTAAAT	ATATACATGT	ATACCTGGAT	TTGCTTGGCT	2820
TCTTTGAGCA	TGTAAGAGAA	ATAAAATTTG	AAAGAATAAT	TTTTCTGTT	CACTGGCTCA	2880
TTTTTTCTTC	AGACACGCAC	TGAACATCTG	TTATTCGGAA	CACCTGGGT	T	

85

Seq ID NO: 162 Protein sequence:  
Protein Accession #: AAA68877.1



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75

1	11	21	31	41	51	
MSLEQRSPHC	KPDEDLEAQG	EDLGLMGAQE	PTGEEEEETS	SSDSKEEEVS	AAGSSSPQOS	60
PQGGASSSIS	VYYTLWSQFD	EGSSSQEEEE	PSSSVDPAPL	EFMFQEALKL	KVAELVHFL	120
HKYRVKEPVT	KAEMLESVIK	NYKRYFPVIF	GKASEFMQVI	FGTDVKEVDP	AGHSYILVTA	180
LGLSCDSMLG	DGHSMPKAAL	LIIVLGVILT	KDNCAPEEVI	WEALSVMGVY	VGKEHMFYGE	240
PRKLLTQDWV	QENYLEYRQV	PGSDPAHYEF	LWGSKAHAET	SYEKVINYL	MLNAREPICY	300
PSLYEEVLGE	EQEGV					

Seq ID NO: 163 DNA sequence  
Nucleic Acid Accession #: AF292100  
Coding sequence: 30-809

15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75

1	11	21	31	41	51	
GGGGGGGGAG	AGGCCTGGAG	GACACCAACA	TGAACAAGTT	GAAATCATCG	CAGAAGGATA	60
AAGTTCGTCA	GTTTATGATC	TTCACACAAT	CTAGTGAAAA	AACAGCAGTA	AGTTGTCTTT	120
CTCAAAATGA	CTGGAAGTTA	GATGTTGCAA	CAGATAATTT	TTTCCAAAAT	CCTGAACTTT	180
ATATACGAGA	GAGTGTAATA	GGATCATTGG	ACAGGAAGAA	GTTAGAACAG	CTGTACAATA	240
GATACAAAGA	CCCTCAAGAT	GAGAATAAAA	TTGGAATAGA	TGGCATAACG	CAGTTCTGTG	300
ATGACCTGGC	ACTCGATCCA	GCCAGCATTG	GTGTGTTGAT	TATTGCGTGG	AAGTTCAGAG	360
CAGCAACACA	GTGCGAGTTC	TCCAAACAGG	AGTTCATGGA	TGGCATGACA	GAATTAGGAT	420
GTGACAGCAT	AGAACAATA	AAGGCCAGAG	TACCCAGAT	GGAACAAGAA	TTGAAAGAAG	480
CAGGACGATT	TAAGGATTTT	TACCAGTTTA	CTTTTAATTT	TGCAAGAAGT	CCAGGACAAA	540
AAGGATTAGA	TCTAGAAATG	GCCATTGCCT	ACTGGAACCT	AGTGCTTAAT	GGAAGATTTA	600
AATTCCTAGA	CTTATGGAAT	AAATTTTGTG	TGGAACATCA	TAAACGATCA	ATACCAAAAG	660
ACACTTGGAA	TCTTCTTTTA	GACTTCAGTA	CGATGATTGC	AGATGACATG	TCTAATTATG	720
ATGAAGAAGG	AGCATGGCCT	GTTCTTATTG	ATGACTTTGT	GGAATTTGCA	CGCCCTCAAA	780
TTGCTGGGAC	AAAAAGTACA	ACAGTGTAGC	ACTAAAGGAA	CCTTTTAGAA	TGTACATAGT	840
CTGTACAATA	AATACAACAG	AAAATTGCAC	AGTCAATTTT	TGCTGGCTGG	ACTGAACTGA	900
AGATCAATCC	TCACAATTCA	GACTGAGGGT	TGAGACAAAA	CTTTAAGGAT	ACATCTTGGA	960
CCATATCGTA	TTTCATTCTT	CTAATGGTGG	TTTGGGCTTG	TCTTCTAGTC	TGGGCCGCTC	1020
TAAACATTTA	TAATGCCAAC	ATTGTGGATT	TCATCTTATA	TCTGTGGACC	ATCCTAGTTT	1080
ATTCTCCCAT	AAGTCTTAGA	AGCTTTATGG	TGATTATTTT	GAGGTTTTCA	TTCTCGCATA	1140
AAGCACAAAT	CTGTCTTCAT	CAGAAAACAG	TTGGCATAAG	AATTAAACAT	ATGAACATCA	1200
CAAAACAATT	TATAAAAAT	TCTTAAATAT	ACGCTTTGGG	CTAGTTGCAA	AGACTATGCT	1260
AATAGCACTT	CCAGTGAGAG	TGATATATTT	AAGTGTACTG	GATCTGGAAT	GGTGTTTTGG	1320
TTTGGGGGGA	ATTTTTTTTT	TTTCTGGGCA	AATCACATAT	GTTGTTGATG	TGAGTATCTG	1380
ATGAAAAAAC	AATGTCAGAA	TAACCGAAT	GAAAATTTT	TAGGATAACT	TGGTGCCTAC	1440
CTGAAAAATG	TATTGTGTTT	TAGACTCTTG	ATTTCAAAAG	GTTCCACAGA	ACTAGTCTGC	1500
GCTTACCTTA	CCCATGTTTA	TATATAGCTG	TCCTACAGGG	AGCTTTTATT	TAGAAAATGT	1560
CTGCATAATG	TTAGATTCTT	CTCCTGTCTA	CATTATGCAC	TACATAATTG	GACTTCATTA	1620
TGCTTTTGAA	ATGCTTATCT	GCCTGTCACA	TAAGTTAAAC	TATTTAATTT	GTTTTGAATG	1680
TTTTGGATTG	CTACACAATA	CAATATTCTA	AATTTAGGCA	TGAGGGTTTT	TTTGTTTTAT	1740
TTTTACTTTT	TTTTTGTCAT	TGCATTATGG	AACACAAATG	AAATCTCTTT	AATTTATAAG	1800
AAGATAGTAG	GAGTTAAATT	TTGAAAATGG	TTGTGATGAG	CCACGAAATT	CAATCTTTAT	1860
AATATAGGTA	CTGCTCTTTC	AGACAAACAG	TCCATTTTTA	ATGACTTCTT	ATTTTGTGTA	1920
AATTACTTTA	ACTGCTAATC	ACTGTGGTTG	CCAAATATTT	ACTTCAGAAG	CAAAGATTTT	1980
CAAACAAGCA	TACACGATGC	AAAATACCAG	TCTGGCTTCT	AGTCTATTTA	CTGTTTGTGT	2040
TCATCTCAGT	TAGCTCAGTT	TTCTCATCAA	AGCAGAAATG	TATCTTGCGT	GTGTGTGTGT	2100
GTGTGTGTGT	GTGTGTGTGT	GTATGTGTGT	ATATATATAT	ATATATATAT	ATATATATAT	2160
TTTTTTTTTT	TTTTTTTTTA	ATTACAAAGT	CCATGAGCTG	CTTTTATGCT	GAAAATGGTC	2220
ATTTCCCTGT	TCATCTACTG	ACATGTGAAG	AAGGGTTTCT	TGCTTTCTTA	AACATTTCCG	2280
TAAGGCAGGC	TAGAAATGTA	ATACTTCAAA	TGTTTGATGA	TTATGGTCTT	TTGATAGGAA	2340
TAGATTCTGC	TTGGGATATA	TATCCAGGCA	CTCTCTAAGG	TCTAGGGTTG	ATATTAACAA	2400
AGGAATGTAC	TTAGAATAGC	AGTACATTTT	ATGCAAATAT	GGAAATTATT	TTAAGAAACA	2460
ATGACATATC	AAAACCTGCT	TTTACATGAT	TTTGAATAG	ACTAGAAAGC	TTTCCCTATA	2520
GACATATTAA	TATTTCAATC	ATAACTTTAA	TTCAAGAATG	CAGTTTACC	AAAAGAAAAA	2580
TTTGAAATTT	TCTATTCAGG	CTACTGGAAT	TGGTTATTAA	AAGAAAAAGG	AAAAGAAGA	2640
ATCTTGCTGC	TTTCAGTATT	TCCTGATTTT	TTTGTAATA	TAAAGAGGAA	CTTCAATTAT	2700
GAAAAATTTT	TAAAGATAT	ATATATCTAT	ATATCTATAT	ATATGTACTG	TTTTGTTTCC	2760
TGTCTTGAAG	ATTTTGAGTT	ATGGTTATTG	GTTTCAGATT	GATTAATTCA	CATATGCTGT	2820
GTTTTCTTTA	AAAGTCATAT	GGGTTCTGTT	CCTAATGCC	TGGATTTTAC	ATATTTTCT	2880
TTTTAAATGC	AAAACCTTTT	CAACAAATA	GTGTTTGTCA	TCAGGTTGGT	ACTAAACATT	2940
TATAATTACT	GTGTAATTAT	AAACAAAAAT	ACATAAAGCT	TTGAATATA	TTATGTAGCA	3000
TAAAAGTTAA	GGTTGTTTCA	TATGATGGCA	TCTTAGAATT	AAACAAACT	TTTACTAGGG	3060
CTGAAAAGAG	AAGACTGATT	TAATGTGGTG	TGATTATTCT	GAAGATAAAT	GTCTGGCTAC	3120
AGGGAATATT	TTGTACTAAA	AAATGATTAC	ACATATGGCT	GTGTGTGTTT	GAGTCTGTGT	3180
CTGTGAGAGA	GCCAGAGAGA	GTGAGAGAGA	TTGACAGAGA	AAGGGAGAGA	CACACACACG	3240
CCCCTTGAAT	TGCTTTAACT	CCTAAGTGGT	TCAGTCCCTA	TTCCGGTAAA	CTCCCCATGC	3300
TGATTTCTTG	TTTAAACTG	AACCATAGGT	ACAGTTTCCT	TTTTGCCAAA	TGTCAAAACA	3360
GGTACAAATT	TTAAAATGTA	ATGCTTTTTA	AATAGAAAAA	TGTATAAAAT	TAGAAGTGCC	3420
CACATATAAA	AAATACTTGA	GATGAAGATT	ATCTTTAGTG	AATATCATCT	GCATATCTCT	3480
GTAAGTTCAA	TTGTGTTTCT	TACAGTCCCT	GTCATATTAC	CAACAGAGGC	AATAAAAGCT	3540
GCAGTGAAAT	TG					

Seq ID NO: 164 Protein sequence:  
Protein Accession #: AAG00606

80  
85

1	11	21	31	41	51	
MNKLKSSQKD	KVRQFMIFTQ	SSEKTAVSCL	SNQDWKLDVA	TDNFFQNPPEL	YIRESVKGSL	60
DRKKLEQLYN	RYKDPQDENK	IGIDGIQQFC	DDLALDPASI	SVLIIAWKFR	AATQCEFSKQ	120
EFMDGMTELG	CDSIEQLKAQ	IPKMEQELKE	PGRFKDFYQF	TFNFAKNPGQ	KGLDLEMAIA	180
YWNLVNLGRF	KFLDLWNKFL	LEHHKRSIPK	DTWNLLLDFFS	TMIADDMSNY	DEEGAWPVLI	240
DDFVEFARPO	IAGTKSTTV					



Seq ID NO: 165 DNA sequence  
Nucleic Acid Accession #: AF256215  
Coding sequence: 220-2028

5	1	11	21	31	41	51	
	CTCCAGTCCG	CATGCTCAGT	AGCTGCTGCC	GGCCGGGCTG	CGGGGCGGCG	TCCGCTGCGC	60
10	GCCTACGGGC	TGCGGTGGCG	GCCGCCGCGG	CACCCGGCAG	GGCCCGCCAG	TCCCCGCTTC	120
	CCTGCTCCAG	AGCCGCCGCC	TGGGCCGCGG	CAGGGCGGGC	CCGGGGCTCC	TCCATGCTGC	180
	CAGCCGCCGG	GCTGCGGAGC	CGACCAAGTG	GCTCCTGCGA	TGGCGGCGGA	AGAGGAGGCT	240
	GCGGCGGGAG	GTAAGTGTT	GAGAGAGGAG	AACCAGTGCA	TTGCTCCTGT	GGTTTCCAGC	300
	CGCGTGAGTC	CAGGGACAAG	ACCAACAGCT	ATGGGGTCTT	TCAGCTCACA	CATGACAGAG	360
15	TTTCCACGAA	AACGCAAAGG	AAGTGATTCA	GACCCATCCC	AAGTGGAAGA	TGGTGAACAC	420
	CAAGTTAAAA	TGAAGGCCTT	CAGAGAAGCT	CATAGCCAAA	CTGAAAAGCG	GAGGAGAGAT	480
	AAAATGAATA	ACCTGATTGA	AGAACTGTCT	GCAATGATCC	CTCAGTGCAA	CCCCATGGCG	540
	CGTAAACTGG	ACAAACTTAC	AGTTTTAAGA	ATGGCTGTTC	AACACTTGAG	ATCTTTAAAA	600
	GGCTTGACAA	ATTCTTATGT	GGGAAGTAAT	TATAGACCAT	CATTTCTTCA	GGATAATGAG	660
20	CTCAGACATT	TAATCCTTAA	GACTGCAGAA	GGCTTCTTAT	TTGTGGTTGG	ATGTGAAAGA	720
	GGAAGAAATTC	TCCTCGTTTC	TAAGTCAGTC	TCCAAAATAC	TTAATTATGA	TCAGGCTAGT	780
	TTGACTGGAC	AAAGCTTATT	TGACTTCTTA	CATCCAAAAG	ATGTTGCCAA	AGTAAAGGAA	840
	CAACTTTCTT	CTTTTGATAT	TCACCAAGA	GAAAAGCTAA	TAGATGCCAA	AACTGGTTTG	900
	CAAGTTTACA	GTAATCTCCA	CGCTGGAAGG	ACACGTGTGT	ATTCTGGCTC	AAGACGATCT	960
25	TTTTTCTGTC	GGATAAAGAG	TTGTAAAATC	TCTGTCAAAG	AAGAGCATGG	ATGCTTACCC	1020
	AACTCAAAGA	AGAAAGAGCA	CAGAAAATTC	TATACTATCC	ATTGCACTGG	TTACTTGAGA	1080
	AGCTGGCCTC	CAATATTGT	TGGAATGAA	GAAGAAAGGA	ACAGTAAGAA	AGACAACAGT	1140
	AATTTTACCT	GCCTTGTGGC	CATTGGAAGA	TTACAGCCAT	ATATTGTTCC	ACAGAACAGT	1200
	GGAGAGATTA	ATGTGAAACC	AACTGAATTT	ATAACCCGGT	TTGCAGTGAA	TGGAAAATTT	1260
30	GTCTATGTAG	ATCAAAGGGC	AACAGCGATT	TTAGGATATC	TGCCTCAGGA	ACTTTTGGGA	1320
	ACTTCTTGTT	ATGAATATTT	TCATCAAGAT	GACCACAATA	ATTTGACTGA	CAAGCACAAA	1380
	GCAGTTCTAC	AGAGTAAGGA	GAAAATACTT	ACAGATTCCCT	ACAAATTCAG	AGCAAAAAGAT	1440
	GGCTCTTTTG	TAACTTTAAA	AAGCCAATGG	TTTAGTTTCA	CAATCCTTGG	GACAAAAGAA	1500
	CTGGAATATA	TTGTATCTGT	CAACACTTTA	GTTTTGGGAC	ATAGTGAGCC	TGGAGAAGCA	1560
35	TCATTTTAC	CTTGTAAGTC	TCAATCATCA	GAAGAATCCT	CTAGACAGTC	CTGTATGAGT	1620
	GTACCTGGAA	TGTCTACTGG	AACAGTACTT	GGTGCTGGTA	GTATTGGAAC	AGATATTGCA	1680
	AATGAAATTC	TGGATTTACA	GAGGTTACAG	TCTTCTTCAT	ACCTTGATGA	TTGAGTCCA	1740
	ACAGGTTTAA	TGAAAGATAC	TCATACTGTA	AACTGCAGGA	GTATGTCAA	TAAGGAGTTG	1800
	TTTCCACCAA	GTCCTTCTGA	AATGGGGGAG	CTAGAGGCTA	CCAGGCAAAA	CCAGAGTACT	1860
40	GTTGCTGTCC	ACAGCCATGA	GCCACTCCTC	ATAGTATGGT	CACAGTTGGA	TTTCGATGCC	1920
	CTATGTGACA	ATGATGACAC	AGCCATGGCT	GCATTTATGA	ATTACTTAGA	AGCAGAGGGG	1980
	GGCCTGGGAG	ACCCTGGGGA	CTTCAGTGAC	ATCCAGTGGA	CCCTCTAGCC	TTTGATTTTT	2040
	AACTCCAAAA	ATGAGAAACA	TTTTAAAGCA	TTATTTACGA	AAAAACTGTC	TCAACTATTC	2100
	TTAAGTACTG	TATTGATATT	GTTTGTATCT	TTTATTAATG	TTCTACCACT	TTTTATAGAT	2160
45	TTGCATCTTC	CTGTACACAG	GATGTGGGGG	AATACGTTTT	CCTCCAAGA	GAACCAAGTT	2220
	TATTATAGAC	TCCTTTATTC	AGTGAATGG	CTTATAATCC	ACTAGTTGCC	ATATTTTTGC	2280
	TAAATATATT	CTAACCAAGA	ATACTACTTA	CATATTGTTT	TGGCTTTGTT	TTATTTTTGA	2340
	TGCAGTTTTT	TTTAGTTGAG	GTAATGTAAT	ATATTGATGT	TTTCCTTTGT	GTCTAAGATT	2400
	GATTTATAAT	AGTAGGTTTG	TATAATTTGG	AACATTTTCC	ATGCCTTGCG	AATTTCTTTA	2460
50	ATTGAGGATA	GGGCTTACAC	ACTTTAAGAA	AACAGTGAGT	ACTTGAACAT	TTAAAGGGAC	2520
	AGTGCAATTT	ATAGTCATAA	TCACATTGAA	TACTGTATTT	GATCTTTGGA	GACTTAGGCA	2580
	AGCACAGAGC	TGGGATATTT	ATGCTCAGTT	GAGCACTTTA	AGATGAATTT	TAAGTGAGAT	2640
	GATTTCTTGC	TTAAAACTCA	GAAAAGTCAA	AGAGTTTCAG	CTTTCCCTTAC	AGAAAAGGAA	2700
	GGATCTTGGG	CCCTAGATCT	TGGGGATTAA	CCTCTGCATA	TAAGATTTAC	TCTTAATAGG	2760
55	CCAGACGTGG	TGCTCACGCC	TGTAATCCCA	GTACTTTGGG	AGGCTGAGAC	GGGCAGATCA	2820
	CTTGAGGTCA	GGAGTTCAAG	ACCAGCCTGG	CCAATATGGT	GAAACCCCGT	TTCTACTAAA	2880
	AATACAAAAA	AAATTACCCA	GGCACTCACT	CTTGAGGTAA	CTAACCAACT	CCCACGATAA	2940
	TGACAGTCCA	TTCATGAGCG	CAAAGGCCTC	ATGACCTAAT	GGCACACACC	TGTAATCCCA	3000
	ACTGCTTGGG	AGGCTGAGGC	GAGAGGATTG	CTTGAACCTG	GGAGGCAGAG	GTTGCAGTGA	3060
60	GCCGAGATCG	CACCACTGCA	CTCCAGTCTG	GGCAACAGAG	TGAGACTTCA	TCTCAAAAAA	3120
	AGTAAAAAAA	AAGATTTAAT	ATAATCACTG	AAGATCTCTA	TTATAGATAG	ATTAGGTTTT	3180
	TGACATTGGA	AACATACTTA	GGGATAGATT	TGTCCTAAAG	GAAAAAAGTA	GGCCCGGGCA	3240
	GATTAAATGT	CTTGTGTAAT	GTCACACATT	AAATTCAGTC	ACACATTAAA	TTCATAGAGT	3300
	TTTAAATGTT	TAATGTATAT	AAACCAGTTT	CTTTATACAC	ATTTGGGAAA	ACATTGGTCT	3360
65	CACGATTAA	ATGATTAACT	AACGTACCTA	GGAAGTATTT	GTAGCTTTCT	AAGTAATTAG	3420
	GCAATTACAG	TTATTGCCCTG	TAACCAAAGG	TAATAAAACA	AAATGACAAG	TACATGTTTA	3480
	AAATTATGAG	GCAATGAGAA	ATAATTTAAA	AACCAATTTT	CTAGTTATAA	TTTAAAATTT	3540
	GGAGAGCATT	TTTAACAGTA	ATTAATCCAG	AGGTGGCTCA	AATTGAGTAT	AAGAATTAAG	3600
	ATTATTTAAA	ATACTGCATG	TCTACCTTCT	CGGGGATCAT	ACTTTATAAC	ACTTTCTGCT	3660
70	TCAGTAGCTC	TTCATAGCTT	GCCAAGTATG	CTCCCATATT	TTCTCTCTCG	TGCCTCGCAA	3720
	ATGAAAGTCA	GATAGGCTGG	GAATCATGG	GGCAGCCCTC	AGACTTCAAT	GTGGGCTTCA	3780
	AATCCAGTTT	CCTGTTCTAT	ATGGTGCTAC	ATCTTTCCAG	AAAAATTTCCC	TCAGAGCCCC	3840
	TCGCCAAAAC	AAAGCATTAT	TTTGACCCTG	CATGCTATTT	CTTTAGCTGT	AGGTGATAGA	3900
75	TTAGAATCTC	TGTCAGACAT	GTTAATGACA	AACATACCAA	CAGACAATAA	CCAAAGCAAA	3960
	TGTTTTCTTC	AAGTGTGAAA	TGTGCAGGGG	CTCGTGGGCA	AGGATGTATT	GGCACACTGT	4020
	CCTCTTGAAC	TGATAGTGTC	CCAGCAATGT	TGGAGGTTGG	CACCATTCCT	GGTCCGACAC	4080
	TTGAGGACCT	GAGAGACATC	AGGTTTAGAA	TAAGCCAAAG	AAATCCTACA	AGATGGGGAG	4140
	AATTGGTGTG	CAGCAGCCTA	AGTGTATAG	TTAAGTCTAA	AGAAGTATGA	AAGATCCCTT	4200
	GTGTTCTCTA	AATTGAGCAG	AGGGGCCTGC	CTACCAATAT	CACTTTTTAG	GGGACTGAAC	4260
80	CATTGCAGGT	TAGACTTGGC	TTCCAAAGAG	TCTGCCTAAG	CCAGGGGTGG	CAGGGTAGGC	4320
	CATCATAGCT	GGATGGCCTC	AAAAGCAGAT	GGGGGCAGAC	TTGCCCTCGT	GATGCCAGGA	4380
	TTTGAGAGGC	AGAGTTTCTA	GAGGGAGACC	AGTGCTGCCCT	CTCACAGTGG	CAGTTTTTTC	4440
	TCTTTGCAAG	AGGAGGGGCT	GTTCAATTCC	ATAGACCAGT	GGGCAGATAG	CCAGTTGAAT	4500
	ACTCTGTGCA	TGGTTTGATC	CTTTATTTAT	TCGCTCTAAT	ATTTTCTGT	AGATCCTTTT	4560
85	GTCTTGACT	CAAAATCTAA	TCCATGCATT	GTATGATACC	GTAGCTCTCC	TAAGGTTTGT	4620
	GTTTCCTTCA	AAATGTTTTA	GTTTTCTTCA	ACTAAATTTG	ATTTTGTCTG	TTAGAAGTGA	4680
	CATATTTTAA	TGGTATACAC	TATGTTCTCT	TTTTCTACTG	CGAGTCAATT	TTTTGAATTT	4740
	TCGTGAGAAA	GAATATATCT	ACAAATTGCA	CGAAAGTATC	ATAAAAAACAG	TACTCTAGAG	4800



	CAGCGCTGTC	CAATAGAAAT	ATAATCTGAG	CCACATGTAT	AATTTTATTT	TCTTCTAGCC	4860
	ACATTAAAGA	AGTAAAAAGA	TACAAGTAGA	ACTAATTTTA	ATGTTTAAAT	TCAGTATATC	4920
	CAAAATATCA	TTTGAACATG	TAATTAATAT	AAAATTATTA	ATGTGATATT	TTACATTCTT	4980
5	TTGGTAATAC	TAGTCTTCAA	AATCTGGTAT	GTATCTTACA	TTGATAGCAC	ATCTCACTTT	5040
	GTACTAGCCA	CATTGCAAGT	GCTCAGTAGC	CACATGTGGC	TAGTGGCTAC	TGCACTGGAC	5100
	AGCACAGTTC	TAGGTTCAC	CCTAACACCC	AAGTCTGTG	GATTAGAATC	CCAGAATCAG	5160
	AGCTGGAAGT	AAACATAGAG	ATCAAACCTC	CTTTTAAAAA	TGAGGACGCT	GAGGCACAGA	5220
	GTTTAAATGG	CTTGCAATGAG	GTCATACAGC	TAAATTCAGC	CTCAACAGGG	TCTTCTGATT	5280
10	CCAGGCACCTC	TTCCCACTCC	ACTACATTAC	TGTAGTGGTA	ATTCTTAGGG	TTAAAAAAG	5340
	TGTAGAGTAG	GCCGGGCGCA	GTGGCTCATG	CCTGTAATCC	CAGCACTTTG	GGAGGCCGAA	5400
	GTGGGCGGAT	CACGAGGTCA	GGAGATCGAG	ACCATCCTGG	CCAACATGGT	GAAACCCCGT	5460
	CTCTACTGAA	AATACAAAGC	AAAATTAGCC	AGGTGTGGTG	GCGGGCGCCT	GTGGTCCCAG	5520
	CTGCTCTGGA	GGCTGAGGCA	GAATGGCGTG	AACCCAGGAG	GCAGAGATGG	CAGTGAGCCA	5580
15	AGATCGCGCC	ACTGCACCCC	AGCCTGGGCG	ACAGAGCGAG	ACTCCATCTC	AAAAAAGAAA	5640
	AAAAAAGAAA	AAGAAAAGAA	AAGAAAAGTC	TAGAGAACAT	TATATTAAGT	GGTTATTATT	5700
	GAAGTAGACC	AAAGTTTATA	CCATAAGGAT	ATTTTTCCTT	AAATACCATG	TTTGAAGAAC	5760
	AATTATTTAT	TGATCCTTGA	ATCTGTAAGA	TCAAATAACA	AGTCTCTATC	CATGTTACCA	5820
	AATTTAACCT	TTTGAATAA	ATAAACTTTA	AAATATCAGA	TGTGTTATTA	CAGGATGATA	5880
20	CTTGGAATCA	AGTGAATGA	GTTATATGGT	CATCACTAAA	TTTAGAAATC	TATTGTGAAA	5940
	CAAAGACAAA	CAGGAAAGTA	CAGAATAGAG	ACTTTTAGTA	AATAAATGGA	ATTTAAAGA	6000
	AAGTGTTTAT	TTACAGTGTC	ACGACAGAAA	AGGATGTCTT	TGTTGTCATA	GTCTTTGAGG	6060
	GATCTCCGTA	AAATCTGGGG	CACAGGTACA	AGAAATAGCC	AATATTTAGT	TCCCAGACCA	6120
	TGTTTAGTAG	TGTCCAGTTT	CAGATCATGC	TGCCAAGAGG	TATCTCCCCC	TCAGGTGGGT	6180
25	CATCACTGAG	CCCTGGAATT	GGAGACTCAT	ACTTGCCAG	CACAATGTTA	CGGGCAGACA	6240
	GGCCGACATC	TATGATTAGC	TAGAAGCCAT	AAAGAAAAGC	TGCTAAGTGG	CCACTAGGTG	6300
	CCACTTTTCT	GTTTTTGTAA	TGCTTTTCTT	AGCAGATCTT	TTTTTTCCAA	GCTCCATGGG	6360
	GCCTATGAGA	GGCATTATG	ATTTTGTGTC	CTACAATAAG	TCAGCTGTGC	TGGTGTGAGT	6420
	TGTTTTATGA	GAAATGCTTT	CCAAGGGAGG	TCTAGGAAGA	TCCTGACACA	TAAGAACTTT	6480
30	GGCTTAGAGA	GCTTTCCAGG	TGTAGTGCCA	ATAAAACTG	ACCTGGAAAG	AAAACCTGCC	6540
	CAGCACGGAA	CATGCTTTCT	GAACCTACTT	GAGAGTGAT	GGTGTATGTC	ACTTCTCATA	6600
	TATTCTTGAG	TTTAGATTG	TCTTTTATAC	AATTTTATAGC	TCTTTTCCAG	TTCACTTGTC	6660
	CTCGTCTGTA	TATTGGTATT	TTTAAATTTT	TGTGGTAAAT	AATGAAAAGA	GTGAAATTAT	6720
35	ATTTTATAAT	TACTCATTTG	TAGTTTTTTT	TTTTAATTTA	ATAAACTTCC	TCCAAAAGT	6780
	GCTCCCTTAA	AA					

Seq ID NO: 166 Protein sequence:  
Protein Accession #: AAG34652

40	1	11	21	31	41	51	
	MAABEEAAG	GKVLREENQC	IAPVVSSRVS	PGTRPTAMGS	FSSHMTEFPR	KRKGSDSDPS	60
	QVEDGEHQVK	MKAFREAHSQ	TEKRRRDKNM	NLIEELSAMI	PQCNPMARKL	DKLTVLRMAV	120
45	QHLRSLKGLT	NSYVGSNYRP	SFLQDNELRH	LILKTAEGFL	FVVGCEGRKI	LFVSKSVSKI	180
	LNVDQASLTG	QSLFDFLHPK	DVAKVKEQLS	SFDISPKEKL	IDAKTGLQVH	SNLHAGRTRV	240
	YSGSRSSFFC	RIKSKISVK	EEHSGCLPNSK	KKEHRKFYTI	HCTGYLRSWP	PNIVGMEEER	300
	NSKKDNSNFT	CLVAIGRLQP	YIVPQNSGEI	NVKPTEFITR	FAVNGKFVYV	DQRATAILGY	360
	LPQELLGTSC	YEYFHQDDHN	NLTDKHKAVL	QSKEKILTDS	YKFRKDGSGF	VTLSQWFSF	420
50	TNPWTKELEY	IVSVNTLVLG	HSEPGAEASFL	PCSSQSSEES	SRQSCMSVPG	MSTGTVLGAG	480
	SIGTDIANEI	LDLQRLQSSS	YLDSSSPTGL	MKDTHTVNCR	SMSNKELEFP	SPSEMGELEA	540
	TRQNQSTVAV	HSHEPLSDG	AQLDFDALCD	NDDTAMAAFM	NYLEAEGGLG	DPGDFSDIQW	600
	TL						

Seq ID NO: 167 DNA sequence  
Nucleic Acid Accession #: NM\_014400  
Coding sequence: 86-1126

60	1	11	21	31	41	51	
	GGTTACTCAT	CCTGGGCTCA	GGTAAGAGGG	CCCCAGCTCG	GAGGCGGCAC	ACCCAGGGGG	60
	GACGCCAAGG	GAGCAGGACG	GAGCCATGGA	CCCCGCCAGG	AAAGCAGGTG	CCCAGGCCAT	120
	GATCTGGACT	GCAGGCTGGC	TGCTGCTGCT	GCTGCTTCGC	GGAGGAGCGC	AGGCCCTGGA	180
	GTGCTACAGC	TGCGTGCAAG	AAGCAGATGA	CGGATGCTCC	CCGAACAAGA	TGAAGACAGT	240
	GAAGTGCGCG	CCGGGCGTGG	ACGTCTGCAC	CGAGGCCGTG	GGGGCGGTGG	AGACCATCCA	300
65	CGGACAATTC	TCGCTGGCAG	TGCSGGGTTG	CGGTTCCGGA	CTCCCCGCGA	AGAATGACCG	360
	CGGCCTGGAT	CTTACGCGGC	TTCTGGCGTT	CATCCAGCTG	CAGCAATGCG	CTCAGGATCG	420
	CTGCAACGCC	AAGCTCAACC	TCACCTCGCG	GGCGCTCGAC	CCGGCAGGTA	ATGAGAGTGC	480
	ATACCCGCCC	AACGGCGTGG	AGTGCTACAG	CTGTGTGGGC	CTGAGCCGGG	AGGCGTGCCA	540
	GGGTACATCG	CCGCCGGTCG	TGAGCTGCTA	CAACGCCAGC	GATCATGTCT	ACAAGGGCTG	600
70	CTTCGACGGC	AACGTCACCT	TGACGGCAGC	TAATGTGACT	GTGTCCTTGC	CTGTCCGGGG	660
	CTGTGTCCAG	GATGAATTCT	GCACTCGGGA	TGGAGTAAAC	GGCCAGGGT	TCACGCTCAG	720
	TGGCTCCTGT	TGCCAGGGGT	CCGCTGTAA	CTCTGACCTC	CGCAACAAGA	CCTACTTCTC	780
	CCCTCGAATC	CCACCCCTTG	TCCGGCTGCC	CCCTCCAGAG	CCCACGACTG	TGGCCTCAAC	840
75	CACATCTGTC	ACCACTTCTA	CCTCGGCCCC	AGTGAGACCC	ACATCCACCA	CCAAACCCAT	900
	GCCAGCGCCA	ACCACTCAGA	CTCCGAGACA	GGGAGTAGAA	CACGAGGCCT	CCCGGGATGA	960
	GGAGCCGAGG	TTGACTGGAG	GCGCCGCTGG	CCACCAGGAC	CGCAGCAATT	CAGGGCAGTA	1020
	TCCTGCAAAA	GGGGGGCCCC	AGCAGCCCCA	TAATAAAGGC	TGTGTGGCTC	CCACAGCTGG	1080
	ATTGGCAGCC	CTTCTGTTGG	CCGTGGCTGC	TGATGTCTTA	CTGTGAGCTT	CTCCACCTGG	1140
80	AAATTTCCCT	CTCACCTACT	TCTCTGGCCC	TGGGTACCCC	TCTTCTCATC	ACTTCTGTGT	1200
	CCCACCACTG	GACTGGGCTG	GCCCAGCCCC	TGTTTTTCCA	ACATTCCCCA	GTATCCCCAG	1260
	CTTCTGCTGC	GCTGGTTTGC	GGCTTTGGGA	AATAAAATAC	CGTTGTATAT	ATTCTGGCAG	1320
	GGGTGTTCTA	GCTTTTGTAG	GACAGCTCCT	GTATCCTTCT	CATCCTTGTC	TCTCCGCTTG	1380
	TCCTCTTGTC	ATGTTAGGAC	AGAGTGAGAG	AAGTCAGCTG	TCACGGGGAA	GGTGAGAGAG	1440
85	AGGATGCTAA	GACTTCTACT	CACCTTCTCC	TAGCCAGCCT	GGACTTTGGA	GCGTGGGGTG	1500
	GGTGGGACAA	TGGCTCCCCA	CTCTAAGCAC	TGCCTCCCTT	ACTCCCCGCA	TCTTTGGGGA	1560
	ATCGGTTCCC	CATATGTCTT	CCTTACTAGA	CTGTGAGCTC	CTCGAGGGCA	GGGACCGTGC	1620
	CTTATGTCTG	TGTGTGATCA	GTTTCTGGCA	CATAAATGCC	TCAATAAAGA	TTTAATTACT	1680



TTGTATAGTG AAAAAAAA

Seq ID NO: 168 Protein sequence:  
Protein Accession #: NP\_055215

1	11	21	31	41	51	
MDPARKAGAQ	AMIWTAGWLL	LLLLLRGGAQA	LECYSCVQKA	DDGCSPNKMK	TVKCAPGVDV	60
CTEAVGAVET	IHGQFSLAVX	GCGSGLPGKN	DRGLDLHGLL	AFIQLQQCAQ	DRCNAKLNLT	120
SRALDPAGNE	SAYPPNGVEC	YSCVGLSREA	CQGTSPFVVS	CYNASDHVYK	GCFDGNVTLT	180
AANVTVSLPV	RGCVQDEFCT	RDGVTGPGFT	LSGSCCQGSR	CNSDLRNKTY	FSPRIPLVR	240
LPPPEPTTVA	STTSVTSTTS	APVRPTSTTK	PMPAPTSQTP	RQGVHEEASR	DEEPRLTGGA	300
AGHQDRSNSG	QYPAKGGPQQ	PHNKGCVAPT	AGLAALLLAV	AAGVLL		

Seq ID NO: 169 DNA sequence  
Nucleic Acid Accession #: NM\_006875  
Coding sequence: 186-1190

1	11	21	31	41	51	
GAATTCGGCA	CGAGCGCGCG	GCGAATCTCA	ACGCTGCGCC	GTCTGCGGGC	GCTTCCGGGC	60
CACCAGTTTC	TCTGCTTTCC	ACCCGCGCGC	CCCCCAGCCC	TGGCTCCCCA	GCTGCGCTGC	120
CCCGGGCGTC	CACGCCCTGC	GGGCTTAGCG	GGTTCACTGG	GCTCAATCTG	CGCAGCGCCA	180
CCTCCATGTT	GACCAAGCCT	CTACAGGGGC	CTCCCGCGCC	CCCCGGGACC	CCCACGCCGC	240
CGCCAGGAGG	CAAGGATCGG	GAAGCGTTCG	AGGCCGAGTA	TGACTCGGC	CCCCTCCTGG	300
GTAAGGGGGG	CTTTGGCACC	GTCTTCGACG	GACACCGCCT	CACAGATCGA	CTCCAGGTGG	360
CCATCAAAGT	GATCCCCCGG	AATCGTGTGC	TGGGCTGGTC	CCCCTTGTC	GACTCAGTCA	420
CATGCCCACT	CGAAGTCGCA	CTGCTATGGA	AAGTGGGTGC	AGGTGGTGGG	CACCCTGGCG	480
TGATCCGCCT	GCTTGACTGG	TTTGAGACAC	AGGAAGGCTT	CATGCTGGTC	CTCGAGCGGC	540
CTTTGCCCGC	CCAGGATCTC	TTTGACTATA	TCACAGAGAA	GGGCCCACTG	GGTGAAGGCC	600
CAAGCCGCTG	CTTCTTTGGC	CAAGTAGTGG	CAGCCATCCA	GCACTGCCAT	TCCCGTGGAG	660
TTGTCCATCG	TGACATCAAG	GATGAGAACA	TCCTGATAGA	CCTACGCCGT	GGCTGTGCCA	720
AACTCATTGA	TTTTGGTTCT	GGTGCCCTGC	TTCATGATGA	ACCCTACACT	GACTTTGATG	780
GGACAAGGGT	GTACAGCCCC	CCAGAGTGGA	TCTCTCGACA	CCAGTACCAT	GCACTCCCGG	840
CCACTGTCTG	GTCACTGGGC	ATCCTCCTCT	ATGACATGGT	GTGTGGGGAC	ATTCCCTTTG	900
AGAGGGACCA	GGAGATTCTG	GAAGCTGAGC	TCCACTTCCC	AGCCCATGTC	TCCCCAGACT	960
GCTGTGCCCT	AATCCGCCCG	TGCCTGGCCC	CCAAACCTTC	TTCCCGACCC	TCACTGGAAG	1020
AGATCCTGCT	GGACCCCTGG	ATGCAAAAC	CAGCCGAGGA	TGTTACCCCT	CAACCCCTCC	1080
AAAGGAGGCC	CTGCCCTTTT	GGCCTGGTCC	TTGCTACCCT	AAGCCTGGCC	TGGCCTGGCC	1140
TGGCCCCCAA	TGGTCAGAA	AGCCATCCCA	TGGCCATGTC	ACAGGGATAG	ATGGACATTT	1200
GTTGACTTGG	TTTTACAGGT	CATTACCAGT	CATTAAAGTC	CAGTATTACT	AAGGTAAGGG	1260
ATTGAGGATC	AGGGGTTAGA	AGACATAAAC	CAAGTTTGCC	CAGTTCCCTT	CCCAATCCTA	1320
CAAAGGAGCC	TTCTCCCGAG	AACCTGTGGT	CCCTGATTTT	GGAGGGGGAA	CTTCTTGCTT	1380
CTCATTTTGC	TAAGGAAGTT	TATTTTGGTG	AAGTTGTTCC	CATTTTGAGC	CCCGGGACTC	1440
TTATTTTGAT	GATGTGTCAC	CCCACATTGG	CACCTCCAC	TACCACCACA	CAAACCTAGT	1500
TCATATGCTT	TTACTTGGGC	AAGGGTGCTT	TCCTTCCAAT	ACCCAGTAG	CTTTTATTTT	1560
AGTAAAGGGA	CCCTTTCCCC	TAGCCTAGGG	TCCCATATTG	GGTCAAGCTG	CTTACCTGCC	1620
TCAGCCCAGG	ATTTTTTATT	TTGGGGGAGG	TAATGCCCTG	TTGTTACCCC	AAGGCTTCTT	1680
TTTTTTTTTT	TTTTTTTTTG	GGTGAGGGGA	CCCTACTTTG	TTATCCCAAG	TGCTCTTATT	1740
CTGGTGAGAA	GAACCTTAAT	TCCATAATTT	GGGAAGGAAT	GGAAGATGGA	CACCACCGGA	1800
CACCACCAGA	CAATAGGATG	GGATGGATGG	TTTTTTGGGG	GATGGGCTAG	GGGAAATAAG	1860
GCTTGCTGTT	TGTTTTCTTG	GGGCGCTCCC	TCCAATTTTG	CAGATTTTTG	CAACCTCCTC	1920
CTGAGCCGGG	ATTGTCCAAT	TACTAAATG	TAAATAATCA	CGTATTGTGG	GGAGGGGAGT	1980
TCCAAGTGTG	CCCTCCTTTT	TTTTCCTGCC	TGGATTATTT	AAAAAGCCAT	GTGTGGAAAC	2040
CCACTATTTA	ATAAAAGTAA	TAGAATCAGA	AAAAAAAAAA	AAAAAAAAAA		

Seq ID NO: 170 Protein sequence:  
Protein Accession #: NP\_006866

1	11	21	31	41	51	
MLTKPLQGGP	APPGTPTPPP	GGKDREAFAE	EYRLGPLLKG	GGFGTVFAGH	RLTDLRLQVAI	60
KVIPRNRVLG	WSPLSDSVTC	PLEVALLWKV	GAGGCHPGVI	RLLDWFETQE	GFMLVLERPL	120
PAQDLFDYIT	EKGPLGEGPS	CFEFGQVVAA	IQHCHSRGVV	HRDIKDENIL	IDLRRGCAKL	180
IDFGSGALLH	DEPYTDFDGT	RVYSPEWIS	RHQYHALPAT	VWSLGILLYD	MVCGDIPFER	240
DQEILEAELH	FPAHVSPDCC	ALIRRCCLAPK	PSSRPSLEEI	LLDPWMQTPA	EDVTPQPLQR	300
RPCPFGLVLA	TLSLAWPGLA	PNGQKSHPMA	MSQG			

Seq ID NO: 171 DNA sequence  
Nucleic Acid Accession #: NM\_003646  
Coding sequence: 89..2875

1	11	21	31	41	51	
GCGGCGCGGA	GCGGGCGTGC	TGAGCCCCGG	CCGCCGGCCC	GGCATGGGCG	TCTCCCGCGG	60
GCCCTCCGCC	GGCCGGGGCT	AGGGCCGGAT	GGAGCCGCGG	GACGGTAGCC	CCGAGGCCCG	120
GAGCAGCGAC	TCCGAGTCGG	CTTCCGCCTC	GTCCAGCGGC	TCCGAGCGCG	ACGCCGGTCC	180
CGAGCCGGAC	AAGGCGCCGC	GGCGACTCAA	CAAGCGGCGC	TTCCCGGGGC	TGCGGCTCTT	240
CGGGCACAGG	AAAGCCATCA	CCAAGTCGGG	CCTCCAGCAC	CTGGCCCCCC	CTCCGCCAC	300
CCCTGGGGCC	CCGTGCAGCG	AGTCAGAGCG	GCAGATCCGG	AGTACAGTGG	ACTGGAGCGA	360
GTCAGCGACA	TATGGGGAGC	ACATCTGGTT	CGAGACCAAC	GTGTCCGGGG	ACTTCTGCTA	420
CGTTGGGGAG	CAGTACTGTG	TAGCCAGGAT	GCTGAAGTCA	GTGTCTCGAA	GAAAGTGCGC	480
AGCCTGCAAG	ATTGTGGTGC	ACACGCCCTG	ACTCGAGCAG	CTGGAGAAGA	TAAATTTCCG	540
CTGTAAGCCG	TCCTTCCGTG	AATCAGGCTC	CAGGAATGTC	CGCGAGCCAA	CCTTTGTACG	600
GCACCACTGG	GTACACAGAC	GACGCCAGGA	CGGCAAGTGT	CGGCACTGTG	GGAAGGGATT	660
CCAGCAGAAG	TTCACCTTCC	ACAGCAAGGA	GATTGTGGCC	ATCAGCTGCT	CGTGGTGCAA	720



	GCAGGCATAC	CACAGCAAGG	TGTCCTGCTT	CATGCTGCAG	CAGATCGAGG	AGCCGTGCTC	780
	GCTGGGGGTC	CACGCAGCCG	TGGTCATCCC	GCCCACCTGG	ATCCTCCGCG	CCCGGAGGCC	840
	CCAGAATACT	CTGAAAGCAA	GCAAGAAGAA	GAAGAGGGCA	TCCTTCAAGA	GGAAGTCCAG	900
5	CAAGAAAGGG	CCTGAGGAGG	GCCGCTGGAG	ACCCTTTCATC	ATCAGGCCCA	CCCCCTCCCC	960
	GCTCATGAAG	CCCCTGCTGG	TGTTTGTGAA	CCCCAAGAGT	GGGGGCAACC	AGGGTGCAAA	1020
	GATCATCCAG	TCTTCTCTCT	GGTATCTCAA	TCCCCGACAA	GTCTTCGACC	TGAGCCAGGG	1080
	AGGGCCCAAG	GAGGCGCTGG	AGATGTACCG	CAAAGTGCAC	AACCTGCGGA	TCCTGGCGTG	1140
	CGGGGGCGAC	GGCACGGTGG	GCTGGATCCT	CTCCACCTTG	GACCAGCTAC	GCCTGAAGCC	1200
10	GCCACCCCTT	GTTGCCATCC	TGCCCCCTGG	TACTGGCAAC	GACTTGGCCC	GAACCCTCAA	1260
	CTGGGGTGGG	GGCTACACAG	ATGAGCCTGT	GTCCAAGATC	CTCTCCCACG	TGGAGGAGGG	1320
	GAACGTGGTA	CAGCTGGACC	GCTGGGACCT	CCACGCTGAG	CCCAACCCCG	AGGCAGGGCC	1380
	TGAGGACCGA	GATGAAGGCG	CCACCGACCG	GTTGCCCTTG	GATGTCTTCA	ACAACCTACTT	1440
	CAGCCTGGGC	TTTGACGCCC	ACGTCACCTT	GGAGTTCCAC	GAGTCTCGAG	AGGCCAACCC	1500
15	AGAGAAATTC	AACAGCCGCT	TTCGGAATAA	GATGTTCTAC	GCCGGGACAG	CTTCTCTCTGA	1560
	CTTCTGTATG	GGCAGCTCCA	AGGACCTGGC	CAAGCACATC	CGAGTGGTGT	GTGATGGAAT	1620
	GGACTTGAAT	CCCAAGATCC	AGGACCTGAA	ACCCAGTGT	GTTGTTTTC	TGAACATCCC	1680
	CAGGTACTGT	GCGGGCACCA	TGCCCTGGGG	CCACCTGGG	GAGCACACG	ACTTTGAGCC	1740
	CCAGCGGCAT	GACGACGGCT	ACCTCGAGGT	CATTGGCTTC	ACCATGACGT	CGTTGGCCGC	1800
20	GCTGCAGGTG	GGCGGACACG	GCGAGCGGCT	GACGCAGTGT	CGCGAGGTGG	TGCTCACCAC	1860
	ATCCAAGGCC	ATCCCGGTGC	AGGTGGATGG	CGAGCCCTGC	AAGCTTGACG	CCTCACGCAT	1920
	CCGCATCGCC	CTGCGCAACC	AGGCCACCAT	GGTGCAAGAG	GCCAAGCGGC	GGAGCGCCGC	1980
	CCCCCTGCAC	AGCGACCAGC	AGCCGGTGCC	AGAGCAGTTG	CGCATCCAGG	TGAGTCGCGT	2040
	CAGCATGCAC	GACTATGAGG	CCCTGCACTA	CGACAAGGAG	CAGCTCAAGG	AGGCCTCTGT	2100
25	GCCGCTGGGC	ACTGTGGTGG	TCCCAGGAGA	CAGTGACCTA	GAGCTCTGCC	GTGCCCACAT	2160
	TGAGAGACTC	CAGCAGGAGC	CCGATGGTGC	TGGAGCCAAG	TCCCCGACAT	GCCAGAAACT	2220
	GTCCCCCAAG	TGGTGCTTCC	TGGACGCCAC	CAGTGCCAGC	CGCTTCTACA	GGATCGACCG	2280
	AGCCCAGGAG	CACCTCAACT	ATGTGACTGA	GATCGCACAG	GATGAGATTT	ATATCCTGGA	2340
	CCCTGAGCTG	CTGGGGGCAT	CGGCCCGGCC	TGACCTCCCA	ACCCCACTT	CCCCCTCTCC	2400
30	CACCTCACCC	TGCTCACCCA	CGCCCCGGTC	ACTGCAAGGG	GATGCTGCAC	CCCCCAAGG	2460
	TGAAGAGCTG	ATTGAGGCTG	CCAAGAGGAA	CGACTTCTGT	AAGCTCCAGG	AGCTGCACCG	2520
	AGCTGGGGGC	GACCTCATGC	ACCGAGACGA	GCAGAGTCGC	ACGCTCCTGC	ACCACGCAGT	2580
	CAGCACTGGC	AGCAAGGATG	TGGTCCGCTA	CCTGCTGGAC	CACGCCCCCC	CAGAGATCCT	2640
	TGATGCGGTG	GAGGAAAACG	GGGAGACCTG	TTTGACCAA	GCAGCGGCC	TGGCCAGCG	2700
35	CACCATCTGC	CATACATCG	TGGAGGCCGG	GGCCTCGCTC	ATGAAGACAG	ACCAGCAGGG	2760
	CGACACTCCC	CGGCAGCGGG	CTGAGAAGGC	TCAGGACACC	GAGCTGGCCG	CCTACCTGGA	2820
	GAACCGGCAG	CATACCAGA	TGATCCAGCG	GGAGGACCAG	GAGACGGCTG	TGTAGCGGGC	2880

Seq ID NO: 172 Protein sequence:  
Protein Accession #: NP\_003637

40	1	11	21	31	41	51	
	MEPRDGSPEA	RSSDSESASA	SSSGSERDAG	PEPDKAPRRL	NKRRFPGLRL	FGHRKAITKS	60
45	GLQHLAPPPP	TPGAPCSESE	RQIRSTVDWS	ESATYGEHIW	FETNVSGDFC	YVGEQYCVAR	120
	MLKSVSRKRC	AACKIVVHTP	CIEQLEKINF	RCKPSFRESG	SRNVREPTFV	RHHWVHRRRQ	180
	DGKSRHCKGK	FQKFTFHSK	EIVAIKSCSW	KQAYHSKVSC	FMLQIEBEP	SLGVHAHVVI	240
	PPTWILRARR	PQNTLKASKK	KKRASFKRKS	SKKGPEEGRW	RPFIIRPTPS	PLMKPLLVFV	300
	NPKSGGNQGA	KIIQSFLWYL	NPRQVFDLSQ	GGPKEALEMY	RKVHNLRLA	CGGDGTVGWI	360
50	LSTLDQLRLK	PPPPVAILPL	GTGNDLARTL	NWGGGYTDEP	VSKILSHVEE	GNVVQLDRWD	420
	LHAEPNPEAG	PEDRDEGATD	RLPLDVFNNY	FSLGFDHVT	LEFHESREAN	PEKFNSRFRN	480
	KMFYAGTAFS	DFLMGSSKDL	AKHIRVVCDD	MDLTPKIQDL	KPQCVVFLNI	PRYCAGTMPW	540
	GHPGEHHDFF	PQRHDDGYLE	VIGFTMTSLA	ALQVGGHGER	LTQCREVVLT	TSKAIPVQVD	600
	GEPCKLAASR	IRIALRNQAT	MVQKAKRRSA	APLHSDQQPV	PEQLRIQVSR	VSMHDYEALH	660
55	YDKEQLKEAS	VPLGTVVVPG	DSDLELCRAH	IERLQQEPDG	AGAKSPTCQK	LSPKWCFLDA	720
	TTASRFYRID	RAQEHLNYVT	EIAQDEIYIL	DPELLGASAR	PDLPTPTSPL	PTSPCSPTPR	780
	SLQGDAAPPQ	GEELIEAAKR	NDFCKLQELH	RAGGDLMHRD	EQSRTLLHHA	VSTGSKDVVR	840
	YLLDHAPPEI	LDAVEENGET	CLHQAAALGQ	RTICHYIVEA	GASLMKTDQQ	GDTPRQRAEK	900
	AQDTELAAYL	ENRQHYQMIQ	REDQETAV				

Seq ID NO: 173 DNA sequence  
Nucleic Acid Accession #: AF232772  
Coding sequence: 1-1662

65	1	11	21	31	41	51	
	ATGCCGGTGC	AGCTGACGAC	AGCCCTGCGT	GTGGTGGGCA	CCAGCCTGTT	TGCCCTGGCA	60
	GTGCTGGGTG	GCATCCTGGC	AGCCTATGTG	ACGGGCTACC	AGTTCATCCA	CACGGAAAAG	120
	CACTACCTGT	CCTTCGGCCT	GTACGGCGCC	ATCCTGGGCC	TGCACCTGCT	CATTGAGAGC	180
70	CTTTTTTGCT	TCCTGGAGCA	CCGGCGCATG	CGACGTGCCG	GCCAGGCCCT	GAAGCTGCCC	240
	TCCCCGCGGC	GGGGCTCGGT	GGCACTGTGC	ATTGCCGCAT	ACCAGGAGGA	CCCTGACTAC	300
	TTGCGCAAGT	GCCTGCGCTC	GGCCCGCGTC	ATCTCCTTCC	CTGACCTCAA	GGTGGTCATG	360
	TGCGTGGATG	GCAACCGCCA	GGAGGACGCC	TACATGCTGG	ACATCTTCCA	CGAGGTGCTG	420
	GGCGGCACCG	AGCAGGCCGG	CTTCTTTGTG	TGGCGCAGCA	ACTTCCATGA	GGCAGGCGAG	480
75	GGTGAGACGG	AGGCCAGCCT	GCAGGAGGGC	ATGGACCGTG	TGCGGGATGT	GGTGGCGGGC	540
	AGCACCTTCT	CGTGCAATCAT	GCAGAAAGTG	GGAGGCAAGC	GCGAGGTCAT	GTACACGGCC	600
	TTCAAGGCCC	TCGGCGATTC	GGTGGACTAC	ATCCAGGTGT	GCGACTCTGA	CACTGTGCTG	660
	GATCCAGCCT	GCACCATCGA	GATGCTTCGA	GTCCTGGAGG	AGGATCCCCA	AGTAGGGGGA	720
	GTGCGGGGAG	ATGTCCAGAT	CCTCAACAAG	TACGACTCAT	GGATTTCCCT	CCTGAGCAGC	780
80	GTGCGGTACT	GGATGGCCTT	CAACGTGGAG	CGGGCCTGCC	AGTCCTACTT	TGGCTGTGTG	840
	CAGTGTATTA	GTGGGCCCTT	GGGCATGTAC	CGCAACAGCC	TCCTCCAGCA	GTTCTCTGGAG	900
	GACTGGTACC	ATCAGAAAGT	CCTAGGCAGC	AAGTGCAGCT	TCGGGGATGA	CCGGCACCTC	960
	ACCAACCGAG	TCCTGAGCCT	TGGCTACCGA	ACTAAGTATA	CCGCGCGCTC	CAAGTGCCCTC	1020
	ACAGAGACCC	CCACTAAGTA	CCTCCGGTGG	CTCAACAGC	AAACCCGCTG	GAGCAAGTCT	1080
85	TACTTCCGGG	AGTGGCTCTA	CAACTCTCTG	TGGTTCCATA	AGCACCACCT	CTGGATGACC	1140
	TACGAGTCAG	TGGTCACGGG	TTTCTTCCCC	TTCTTCTCTA	TTGCCACGGT	TATACAGCTT	1200
	TTCTACCGGG	GCCGCATCTG	GAACATTCTC	CTCTTCTCTG	TGACGGTGCA	GCTGGTGGGC	1260
	ATTATCAAGG	CCACCTACGC	CTGCTTCTCT	CGGGGCAATG	CAGAGATGAT	CTTCATGTCC	1320



	CTCTACTCCC	TCCTCTATAT	GTCCAGCCTT	CTGCCGGCCA	AGATCTTTGC	CATTGCTACC	1380
	ATCAACAAAT	CTGGCTGGGG	CACCTCTGGC	CGAAAAACCA	TTGTGGTGAA	CTTCATTGGC	1440
	CTCATTCCCTG	TGTCCATCTG	GGTGGCAGTT	CTCCTGGAGG	GGCTGGCCTA	CACAGCTTAT	1500
5	TGCCAGGACC	TGTTCACTGA	GACAGAGCTA	GCCTTCCTTG	TCTCTGGGGC	TATACTGTAT	1560
	GGCTGCTACT	GGGTGGCCCT	CCTCATGCTA	TATCTGGCCA	TCATCGCCCC	GCGATGTGGG	1620
	AAGAAGCCGG	AGCAGTACAG	CTTGGCTTTT	GCTGAGGTGT	GACATGGCCC	CCAAGCAGAG	1680
	CGGGTAAAGT	GCAATGGGTA	AGGGAGGGAA	GGGGAATGGA	AGAGAAAAGA	CAGGGTGGGA	1740
	GGGAGGAGGG	AGTGCCTGTG	TTTAGTCTCT	TAATGGTCCA	AAGGACAAAT	CTAAAATGCA	1800
10	AAGAACGGTG	ATGTAGTATG	GCCTGACAGC	TCTGTTTAGA	GGAGGCAACA	CTGATCCCCC	1860
	AGATGCAGGG	CTGCAGGGGA	TTCTGTGTTT	TCAGACTGCC	TGTCTGCTTG	CATCTGCACA	1920
	TAGGCAGTAG	CCTCCTCCTG	GGCTCCAGAG	GGCACTCAGA	AGTTGTGCTA	AACCAAGTTA	1980
	AGTCCCATTC	AGTGGCAACT	TGTGATAGGT	ACCTGAGTGA	CGGCAACCTG	CGGAAGGAGG	2040
	TTCTCCCAGC	CCATCTGAAC	ACAACCAGAG	GTGGCAGGAG	AATTTCTACT	GAGCGAGGTG	2100
	GGCCGGTTAG	TGTATGTCAC	CCCCACCCCA	CCCATAAGTA	GTCATCAATG	CAATAAGATT	2160
15	GCGCTGAGA	TACAAGGCC	AGAAGCCTGA	TCTTTGGGCA	TCAGAAAACA	GGGTCCAGGA	2220
	ATGGTGCTTT	ATGTGAGATA	CCCCACTCCA	CATCAACATT	CCAGGGATGA	GCCAAACCAG	2280
	CAGGGAGTTA	GCACTGAACT	GCTTTTAAAA	GTGCACATTA	AAAAGGAAAG	TTTGCCAGGA	2340
	GGAACAAAGA	GATTGTGGTG	GTGCTAAAGG	AGGCCATAAG	CTACACAGAG	GCCTTGGGTG	2400
20	TTCCACCTGG	AAACTGCTCA	GACGCTAGA	TGGGTCTTTA	GCTTGTCTGT	GATCTCTGCT	2460
	GGGGAGATAA	AAAGATTAAG	CCCCAACATG	TTCAGAAAAG	AAGTGAAGTC	TTGGGTATTT	2520
	TAACTGTAT	ACTCTTGAAT	TCCTCTCAAA	TTGAGCTCTG	ATCTGAGGCT	AAGACACACT	2580
	CCCCACTTCA	CTTTCTTCAA	AGCCACATTT	TTTGAGGTAT	CACTGCAGTC	ACCTCTTCTA	2640
	CCCTCATCAT	CATAGGTAAG	GTTTTCAAGG	TGGCAATTGG	GGCGGAGCCC	CGGCTTCTTA	2700
25	TAGAAGCTTC	AGCAGGAGGC	AAGCGTGTTT	TCAGCACATA	TGGGAACAT	GAGGAGCCTC	2760
	TGATCAAATT	GGCTACAATC	TTGGAGCTGC	TTGGACGGAT	TCCTTGCCAG	CCGGGTAGC	2820
	ATGTGTGACT	TTCAGGCTAC	TGTTCTTGAC	AATCATCTCC	AATGGAAAGC	TTTTCACTGT	2880
	TCCCAAAGTG	AACCTCTCAA	TCCAAAATGG	TTATCTTTGA	GACCATCCAT	TCTCCTCAGT	2940
	GGCTTCTCCA	GGGAATTCTT	ACAGCCAAGT	TGTGACAGTC	ACTGCATTTG	CCTGCTTCTT	3000
30	TCCAGAAACC	AAACTAGGAG	ATGAAACTGG	TTCTACATC	CTAAGGTTCT	TGCTTTCTCT	3060
	CTCATGCCTC	CTGAGGCTGT	TTTTGGCTGT	TTTCCCTCTG	CTGCTTTTGG	GGAATGAGGG	3120
	GAAGCCATTT	TCCAAGTGAC	TTGCAATCCA	GGCTGTTCTC	AGCGTTTGA	GTTTAAACC	3180
	TGGGATCCTG	ACTAAGCCTT	TGACTTAAGG	GTTGCTTGCT	TGCCCTCCAA	ATGTCCTTTC	3240
	TCAAAGGGGC	CAACTAACC	GTGCAGAACC	AGCACTAAGG	TGGACAGCAG	ACAAGAGGGC	3300
35	AAGCCTCTAA	TGTACCAAGT	GCTTCCCTACA	AAGACGCAAG	GTGTGCTCCG	AACCACAGAT	3360
	GGGCAAAACC	TGGTGCTTTC	CTTCATCTCC	CACGAACTCA	AGGGTTTTC	AAGTGTAGCT	3420
	AACAGTTGCC	ACATCACACA	GACCTCCAGT	TTCTGGTAAG	ACTGCTGGTT	GACATCAGAC	3480
	CCAACCCATT	GAAGGCTGGA	AGGCAGCAGG	CATTTGCTAA	GGCAGCTGAT	CCAGGCAATC	3540
	GTTCTGCTGG	CCAAGAAGTT	AAACTATTTT	GAGCATTAGA	ATGGAGGAAA	TCCGGTCAGC	3600
40	CAAGTGACAG	GTTTCAGACT	CGCTAAGGGC	TTGTTTTTCT	TCAGCATTTA	CTTGAAGATT	3660
	AATGTAGGAT	GACAGGCTCT	CTGGCTGTC	CTACCATCAG	CTCTGCCTTG	CACTGTGGTC	3720
	GTCACCTTTC	CTCAAAATCAA	AAACAGGTCG	GTACAGGTAG	TGGGCTCACA	ACGTTTGACC	3780
	TCGACTGGTT	TTTCTAAGTT	ATTTTGTACA	TTTTTCAGCA	GCAAAACCAA	ACTGGGTCTT	3840
	CAGCTTTATC	CCCGTTTCTT	GCAAGGGAAG	AGCCTTTATA	CAATTGGACG	CATTTTGGTT	3900
45	TTTCTCTATT	GAGAATTCAA	ATCCTCTTTT	GTATTGTTTT	TACAATAATT	TGTAAACATA	3960
	TTTATTTTAA	CCTGCTTTTT	TTTTTTTTTT	TAATTTTCAG	GTCAAGTTTT	TTTACTTGCA	4020
	CTTATTTGTC	AAAATAAAGA	TTCTCACAT				

Seq ID NO: 174 Protein sequence:  
Protein Accession #: AAF36984

50	1	11	21	31	41	51	
	MPVQLTTALR	VVGTSFLFALA	VLGGILAAVY	TGYQFIHTEK	HYLSFGLYGA	ILGLHLLIQS	60
55	LFAPLEHRRM	RRAGQALKLP	SPRRGSVALC	IAAYQEDPDY	LRKCLRSAQR	ISFPDLKVVM	120
	VVDGNRQEDA	YMLDIFHEVL	GGTEQAGFFV	WRSNFHEAGE	GETEASLQEG	MDRVRDVVRA	180
	STFSCIMQKW	GGKREVMTYA	FKALGDSVDY	IQVCDSDTVL	DPACTIEMLR	VLEEDPQVGG	240
	VGGDVQILNK	YDSWISFLSS	VRYWMAFNVE	RACQSYFGCV	QCISGPLGMY	RNSLLQQFLE	300
	DWYHQKFLGS	KCSFGDDRHL	TNRVLSLGYR	TKYTARSKCL	TETPTKYLRW	LNQQTRWSKS	360
60	YFREWLNSL	WFHKHHLWMT	YESVVTGFFP	PFLIATVIQL	FYRGRIWNIL	LFLLTVQLVG	420
	IIKATYACFL	RGNAEMIFMS	LYSLLYMSL	LPAKIFAIAI	INKSGWGTSG	RKTIVVNFIS	480
	LIPVSIWVAV	LLEGLAYTAY	CQDLFSETEL	AFLVSGAILY	GCYWVALLML	YLAIIRRCG	540
	KKPEQYSLAF	AEV					

Seq ID NO: 175 DNA sequence  
Nucleic Acid Accession #: NM\_000691  
Coding sequence: 43..1404

70	1	11	21	31	41	51	
	CCAGGAGCCC	CAGTTACCGG	GAGAGGCTGT	GTCAAAGGCG	CCATGAGCAA	GATCAGCGAG	60
	GCCGTGAAGC	GCGCCCGCGC	CGCCTTCAGC	TCGGGCAGGA	CCCGTCCGCT	GCAGTTCCGA	120
	TTCCAGCAGC	TGGAGGCGCT	GCAGCGCCTG	ATCCAGGAGC	AGGAGCAGGA	GCTGGTGGGC	180
75	GCGCTGGCCG	CAGACCTGCA	CAAGAATGAA	TGGAACGCCCT	ACTATGAGGA	GGTGGTGTAC	240
	GTCCTAGAGG	AGATCGAGTA	CATGATCCAG	AAGCTCCCTG	AGTGGGCCCG	GGATGAGCCC	300
	GTGGAGAAGA	CGCCCCAGAC	TCAGCAGGAC	GAGCTCTACA	TCCACTCGGA	GCCACTGGGC	360
	GTGGTCTCTG	TCATTGGCAC	CTGGAACCTAC	CCCTTCAACC	TCACCATCCA	GCCCATGGTG	420
	GGCGCCATCG	CTGCAGGGAA	CGCAGTGFTC	CTCAAGCCCT	CGGAGCTGAG	TGAGAACATG	480
80	GCGAGCCTGC	TGGCTACCAT	CATCCCCCAG	TACCTGGACA	AGGATCTGTA	CCCAGTAATC	540
	AATGGGGGTG	TCCCTGAGAC	CACGGAGCTG	CTCAAGGAGA	GGTTCGACCA	TATCCTGTAC	600
	ACGGGCAGCA	CGGGGGTGGG	GAAGATCATC	ATGACGGCTG	CTGCCAAGCA	CCTGACCCCT	660
	GTCACGCTGG	AGCTGGGAGG	GAAGAGTCCC	TGCTACGTGG	ACAAGAACTG	TGACCTGGAC	720
	GTGGCCTGCC	GAGGCATCGC	CTGGGGGAAA	TTTATGAACA	GTGGCCAGAC	CTGCGTGGCC	780
85	CCAGACTACA	TCCTCTGTGA	CCCCTCGATC	CAGAACCAAA	TTGTGGAGAA	GCTCAAGAAG	840
	TCACTGAAAG	AGTTCTACGG	GGAAGATGCT	AAGAAATCCC	GGGACTATGG	AAGAATCATT	900
	AGTGCCCGGC	ACTTCCAGAG	GGTGATGGGC	CTGATTGAGG	GCCAGAAGGT	GGCTTATGGG	960
	GGCACCAGGG	ATGCCGCCAC	TCGCTACATA	GCCCCACCA	TCCTCACGGA	CGTGGACCCC	1020



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

CAGTCCCCGG	TGATGCAAGA	GGAGATCTTC	GGGCCTGTGC	TGCCCCATCGT	GTGCGTGCGC	1080
AGCCTGGAGG	AGGCCATCCA	GTTTCATCAAC	CAGCGTGAGA	AGCCCCTGGC	CCTCTACATG	1140
TTCTCCAGCA	ACGACAAGGT	GATTAAGAAG	ATGATTGCAG	AGACATCCAG	TGGTGGGGTG	1200
GCGGCCAACG	ATGTCATCGT	CCACATCACC	TTGCACTCTC	TGCCCCTCGG	GGGCGTGGGG	1260
AACAGCGGCA	TGGGATCCTA	CCATGGCAAG	AAGAGCTTCG	AGACTTTCTC	TCACCGCCGC	1320
TCTTGCCTGG	TGAGGCCTCT	GATGAATGAT	GAAGGCCTGA	AGGTCAGATA	CCCCCGGAGC	1380
CCGGCCAAGA	TGACCCAGCA	CTGAGGAGGG	GTTGCTCCGC	CTGGCCTGGC	CATACTGTGT	1440
CCCATCGGAG	TGCGGACCAC	CCTCACTGGC	TCTCCTGGCC	CTGGAGAATC	GCTCCTGCAG	1500
CCCCAGCCCA	GCCCCACTCC	TCTGCTGACC	TGCTGACCTG	TGCACACCCC	ACTCCACAT	1560
GGGCCAGGC	CTCACCATT	CAAGTCTCCA	CCCCTTTCTA	GACCAATAAA	GAGACAAATA	1620
CAATTTTCTA	ACTCGG					

Seq ID NO: 176 Protein sequence:  
Protein Accession #: NP\_000682

15  
20  
25

1	11	21	31	41	51	
MSKISEAVKR	ARAAFSSGRT	RPLQFRFQQL	EALQRLIQEQ	EQELVGALAA	DLHKNEWNAY	60
YEEVVYVLEE	IEYMIQKLPE	WAADEPVEKT	PQTQQDELYI	HSEPLGVVLV	IGTWNYPFNL	120
TIQPMVGAIA	AGNAVVLKPS	ELSENMASLL	ATIIPQYLDK	DLYPVINGGV	PETTELLKER	180
FDHILYTGST	GVGKIIMTAA	AKHLTPVTLE	LGGKSPCYVD	KNCDLDVACR	RIAWGKFMNS	240
GQTCVAPDYI	LCDPSIQNQI	VEKLKSLKE	FYGEDAKKSR	DYGRIISARH	FQRMVGLIEG	300
QKVAYGGTGD	AATRYIAPTI	LTDVDPQSPV	MQEEIFGPVL	PIVCVRSLEE	AIQFINQREK	360
PLALYMFSSN	DKVIKKMIAE	TSSGGVAAND	VIVHITLHSL	PFGGVGNSGM	GSYHGKKSFE	420
TFSHRRSCLV	RPLMNDGLK	VRYPSPAKM	TQH			

Seq ID NO: 177 DNA sequence  
Nucleic Acid Accession #: NM\_001067.1  
Coding sequence: 108-4703

30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
CTAACCGACG	CGCGTCTGTG	GAGAAGCGGC	TTGGTTCGGGG	GTGGTCTCGT	GGGGTCCTGC	60
CTGTTTAGTC	GCTTTCAGGG	TTCTTGAGCC	CCTTCACGAC	CGTCACCATG	GAAGTGTCAC	120
CATTGCAGCC	TGTAAATGAA	AATATGCAAG	TCAACAAAAT	AAAGAAAAAT	GAAGATGCTA	180
AGAAAAGACT	GTCTGTTGAA	AGAATCTATC	AAAAGAAAAC	ACAATTGGAA	CATATTTTGC	240
TCCGCCCAGA	CACCTACATT	GGTTCGTGTT	AATTAGTGAC	CCAGCAAATG	TGGGTTTACG	300
ATGAAGATGT	TGGCATTAAAC	TATAGGGAAG	TCACTTTTGT	TCCTGGTTTG	TACAAAATCT	360
TTGATGAGAT	TCTAGTTAAT	GCTGCGGACA	ACAAACAAAG	GGACCCAAA	ATGTCCTGTA	420
TTAGAGTCAC	AATTGATCCG	GAAAACAATT	TAATTAGTAT	ATGGAATAAT	GGAAAAGGTA	480
TTCTGTGTTG	TGAACACAAA	GTTGAAAAGA	TGTATGTCCC	AGCTCTCATA	TTTGACACGC	540
TCCTAACTTC	TAGTAACTAT	GATGATGATG	AAAAGAAAGT	GACAGGTGGT	CGAAATGGCT	600
ATGGAGCCAA	ATTGTGTAAC	ATATTCAGTA	CCAAATTTAC	TGTGGAAACA	GCCAGTAGAG	660
AATACAAGAA	AATGTTCAAA	CAGACATGGA	TGGATAATAT	GGGAAGAGCT	GGTGAGATGG	720
AACTCAAGCC	CTTCAATGGA	GAAGATTATA	CATGTATCAC	CTTTCAGCCT	GATTTGTCTA	780
AGTTTAAAT	GCAAAGCCTG	GACAAAGATA	TTGTGTCAC	AATGGTCAGA	AGAGCATATG	840
ATATTGCTGG	ATCCACCAA	GATGTCAAAG	TCTTTCTTAA	TGGAAATAAA	CTGCCAGTAA	900
AAGGATTTCT	TAGTTATGTG	GACATGTATT	TGAAGGACAA	GTTGGATGAA	ACTGGTAACT	960
CCTTGAAAGT	AATACATGAA	CAAGTAAACC	ACAGGTGGGA	AGTGTGTTTA	ACTATGAGTG	1020
AAAAAGGCTT	TCAGCAAATT	AGCTTTGTCA	ACAGCATTGC	TACATCCAAG	GGTGGCAGAC	1080
ATGTTGATTA	TGTAGCTGAT	CAGATTGTGA	CTAAACTTGT	TGATGTTGTG	AAGAAGAAGA	1140
ACAAGGGTGG	TGTTGCAGTA	AAAGCACATC	AGGTGAAAAA	TCACATGTGG	ATTTTGTGTA	1200
ATGCCTTAAT	TGAAAACCCA	ACCTTTGACT	CTCAGACAAA	AGAAAACATG	ACTTTACAAC	1260
CCAAGAGCTT	TGGATCAACA	TGCCAATTGA	GTGAAAAAAT	TATCAAAGCT	GCCATTGGCT	1320
GTGGTATTGT	AGAAAGCATA	CTAAACTGGG	TGAAGTTTAA	GGCCCAAGTC	CAGTTAAACA	1380
AGAAGTGTTT	AGCTGTAAAA	CATAATAGAA	TCAAGGGAAT	TCCCAAATCT	GATGATGCCA	1440
ATGATGCAGG	GGGCCGAAAC	TCCACTGAGT	GTACGCTTAT	CCTGACTGAG	GGAGATTCAG	1500
CCAAAACCTT	GGCTGTTTCA	GGCCTTGGTG	TGGTTGGGAG	AGACAAATAT	GGGGTTTTC	1560
CTCTTAGAGG	AAAAATACTC	AATGTTTCGAG	AAGCTTCTCA	TAAGCAGATC	ATGGAAAATG	1620
CTGAGATTAA	CAATATCATC	AAGATTGTGG	GCTTTCAGTA	CAAGAAAAAC	TATGAAGATG	1680
AAGATTCATT	GAAGACGCTT	CGTTATGGGA	AGATAATGAT	TATGACAGAT	CAGGACCAAG	1740
ATGGTTCCCA	CATCAAAGGC	TTGCTGATTA	ATTTTATCCA	TCACAATCTG	CCCTCTCTTC	1800
TGCGACATCG	TTTTCTGGAG	GAATTTATCA	CTCCCATTTG	AAAGGTATCT	AAAAACAAGC	1860
AAGAAATGGC	ATTTTACAGC	CTTCTGAAT	TTGAAGAGTG	GAAGAGTTCT	ACTCCAAATC	1920
ATAAAAAATG	GAAAGTCAAA	TATTACAAAG	GTTTGGGCAC	CAGCACATCA	AAGGAAGCTA	1980
AAGAATACTT	TGCAGATATG	AAAAGACATC	GTATCCAGTT	CAAATATTCT	GGTCTGGAAG	2040
ATGATGCTGC	TATCAGCCTG	GCCTTTAGCA	AAAAACAGAT	AGATGATCGA	AAGGAATGGT	2100
TAACTAATTT	CATGGAGGAT	AGAAGACAAC	GAAAGTTACT	TGGGCTTCCT	GAGGATTACT	2160
TGTATGGACA	AACTACCACA	TATCTGACAT	ATAATGACTT	CATCAACAAG	GAACCTATCT	2220
TGTTCTCAAA	TTCTGATAAC	GAGAGATCTA	TCCCTTCTAT	GGTGGATGGT	TTGAAACCAG	2280
GTCAGAGAAA	GGTTTGTGTT	ACTTGCTTCA	AACGGAATGA	CAAGCGAGAA	GTAAAGGTTG	2340
CCCAATTAGC	TGGATCAGTG	GCTGAAATGT	CTTCTTATCA	TCATGGTGAG	ATGTCATAAA	2400
TGATGACCAT	TATCAATTTG	GCTCAGAATT	TTGTGGGTAG	CAATAATCTA	AACCTCTTGC	2460
AGCCCATTTG	TCAGTTTGGT	ACCAGGCTAC	ATGGTGGCAA	GGATTCTGCT	AGTCCACGAT	2520
ACATCTTTAC	AATGCTCAGC	TCTTTGGCTC	GATTGTTATT	TCCACCAAAA	GATGATCACA	2580
CGTTGAAGTT	TTTATATGAT	GACAACCAGC	GTGTTGAGCC	TGAATGGTAC	ATTCTTATTA	2640
TTCCCATGGT	GCTGATAAAT	GGTGCTGAAG	GAATCGGTAC	TGGGTGGTCC	TGCAAAATCC	2700
CCAACTTTGA	TGTGCGTGAA	ATTGTAAAT	ACATCAGGCG	TTTGATGGAT	GGAGAAGAAC	2760
CTTTGCCAAT	GCTTCCAAGT	TACAAGAACT	TCAAGGGTAC	TATTGAAGAA	CTGGCTCCAA	2820
ATCAATATGT	GATTAGTGGT	GAAGTAGCTA	TTCTTAATTC	TACAACCATT	GAAATCTCAG	2880
AGCTTCCCGT	CAGAACATGG	ACCCAGACAT	ACAAAGAACA	AGTTCTAGAA	CCCATGTTGA	2940
ATGGCACCGA	GAAGACACCT	CCTCTCATAA	CAGACTATAG	GGAATACCAT	ACAGATACCA	3000
CTGTGAAATT	TGTTGTGAAG	ATGACTGAAG	AAAAACTGGC	AGAGGCAGAG	AGAGTTGGAC	3060
TACACAAAGT	CTTCAAACTC	CAAACTAGTC	TCACATGCAA	CTCTATGGTG	CTTTTGTGAC	3120
ACGTAGGCTG	TTTAAAGAAA	TAAAGACCGG	TGTTGGATAT	TCTAAGAGAC	TTTTTTGAAC	3180
TCAGACTTAA	ATATTATGGA	TTAAGAAAAG	AATGGCTCCT	AGGAATGCTT	GGTGCTGAAT	3240
CTGCTAAACT	GAATAATCAG	GCTCGCTTTA	TCTTAGAGAA	AATAGATGGC	AAAATAATCA	3300



	TTGAAATAA	GCCTAAGAAA	GAATTAATTA	AAGTTCTGAT	TCAGAGGGGA	TATGATTCCG	3360
	ATCCTGTGAA	GGCCTGGAAA	GAAGCCCAGC	AAAAGGTTCC	AGATGAAGAA	GAAAATGAAG	3420
	AGAGTGACAA	CGAAAAGGAA	ACTGAAAAGA	GTGACTCCGT	AACAGATTCT	GGACCAACCT	3480
5	TCAACTATCT	TCTTGATATG	CCCCTTTGGT	ATTTAACCAG	GGAAAAGAAA	GATGAACTCT	3540
	GCAGGCTAAG	AAATGAAAAA	GAACAAGAGC	TGGACACATT	AAAAAGAAAAG	AGTCCATCAG	3600
	ATTTGTGGAA	AGAAGACTTG	GCTACATTTA	TTGAAGAATT	GGAGGCTGTT	GAAGCCAAGG	3660
	AAAAACAAGA	TGAACAAGTC	GGACTTCCTG	GGAAAGGGGG	GAAGGCCAAG	GGGAAAAAAA	3720
	CACAAATGGC	TGAAGTTTTG	CCTTCTCCGC	GTGGTCAAAG	AGTCATTCCA	CGAATAACCA	3780
10	TAGAAATGAA	AGCAGAGGCA	GAAAAGAAAA	ATAAAAAGAA	AATTAAGAAT	GAAAATACTG	3840
	AAGGAAGCCC	TCAAGAAGAT	GGTGTGGAAC	TAGAAGGCCCT	AAAACAAAGA	TTAGAAAAGA	3900
	AACAGAAAAG	AGAACCAGGT	ACAAAGACAA	AGAAACAAAC	TACATTGGCA	TTTAAGCCAA	3960
	TCAAAAAAGG	AAAGAAGAGA	AATCCCTGGC	CTGATTGAGA	ATCAGATAGG	AGCAGTGACG	4020
	AAAGTAATTT	TGATGTCCCT	CCACGAGAAA	CAGAGCCACG	GAGAGCAGCA	ACAAAAACAA	4080
	AATTACACAAT	GGATTGGAT	TCAGATGAAG	ATTTCTCAGA	TTTTGATGAA	AAAACGTATG	4140
15	ATGAAGATTT	TGTCCCATCA	GATGCTAGTC	CACCTAAGAC	CAAAACTTCC	CCAAAACTTA	4200
	GTAACAAAGA	ACTGAAACCA	CAGAAAAGTG	TCGTGTCAGA	CCTTGAAGCT	GATGATGTTA	4260
	AGGGCAGTGT	ACCACTGTCT	TCAAGCCCTC	CTGCTACACA	TTTCCCAGAT	GAAACTGAAA	4320
	TTACAAACCC	AGTTCCTAAA	AAGAATGTGA	CAGTGAAGAA	GACAGCAGCA	AAAAGTCAGT	4380
20	CTTCCACCTC	CACTACCGGT	GCCAAAAAAA	GGGCTGCCCC	AAAAGGAACT	AAAAGGGATC	4440
	CAGCTTTGAA	TTCTGGTGTG	TCTCAAAAGC	CTGATCCTGC	CAAAACCAAG	AATCGCCGCA	4500
	AAAGGAAGCC	ATCCACTTCT	GATGATTCTG	ACTCTAATTT	TGAGAAAAAT	GTTCGAAAAG	4560
	CAGTCACAAG	CAAGAAATCC	AAGGGGGAGA	GTGATGACTT	CCATATGGAC	TTTGACTCAG	4620
	CTGTGGCTCC	TCGGGCAAAA	TCTGTACGGG	CAAAGAAACC	TATAAAGTAC	CTGGAAGAGT	4680
25	CAGATGAAGA	TGATCTGTTT	TAAAATGTGA	GGCGATTATT	TTAAGTAATT	ATCTTACCAA	4740
	GCCCAAGACT	GGTTTTTAAAG	TTACCTGAAG	CTCTTAACTT	CCTCCCCTCT	GAATTTAGTT	4800
	TGGGGAAGGT	GTTTTTTAGTA	CAAGACATCA	AAGTGAAGTA	AAGCCCAAGT	GTTCTTTAGC	4860
	TTTTTATAAT	ACTGTCTAAA	TAGTGACCAT	CTCATGGGCA	TTGTTTTCTT	CTCTGCTTTG	4920
	TCTGTGTTTT	GAGTCTGCTT	TCTTTTGTCT	TTAAACCTG	ATTTTTAAGT	TCTTCTGAAC	4980
30	TGTAGAAATA	GCTATCTGAT	CACTTCAGCG	TAAAGCAGTG	TGTTTATTAA	CCATCCACTA	5040
	AGCTAAAACT	AGAGCAGTTT	GATTTAAAAG	TGTCACTCTT	CCTCCTTTTC	TACTTTCAGT	5100
	AGATATGAGA	TAGAGCATAA	TTATCTGTTT	TATCTTAGTT	TTATACATAA	TTTACCATCA	5160
	GATAGAAGTT	TATGGTTCTA	GTACAGATAC	TCTACTACAC	TCAGCCTCTT	ATGTGCCAAG	5220
	TTTTTCTTTA	AGCAATGAGA	AATTGCTCAT	TCTCTTCATC	TTCTCAAATC	ATCAGAGGCC	5280
35	AAAGAAAAAC	ACTTTGGCTG	TGTCTATAAC	TTGACACAGT	CAATAGAATG	AAGAAAATTA	5340
	GAGTAGTTAT	GTGATTATTT	CAGCTCTTGA	CCTGTCCCCT	CTGGCTGCCT	CTGAGTCTGA	5400
	ATCTCCCAA	GAGAGAAACC	AATTCTAAG	AGGACTGGAT	TGCAGAAAGC	TCGGGGACAA	5460
	CATTTGATCC	AAGATCTTAA	ATGTTATATT	GATAACCATG	CTCAGCAATG	AGCTATTAGA	5520
	TTCATTTTGG	GAAATCTCCA	TAATTTCAAT	TTGTAAACTT	TGTTAAGACC	TGTCTACATT	5580
40	GTTATATGTG	TGTGACTTGA	GTAATGTTAT	CAACGTTTTT	GTAAATATTT	ACTATGTTTT	5640
	TCTATTAGCT	AAATTCCAAC	AATTTTGTAC	TTTAATAAAA	TGTTCTAAAC	ATTGTC	

Seq ID NO: 178 Protein sequence:  
Protein Accession #: NP\_001058.1

45	1	11	21	31	41	51	
	MEVSPLPQVN	ENMQVNKIKK	NEDAKKRLSV	ERIQKKTQL	EHILLRPDTY	IGSVELVTQQ	60
	MWVYDEDVGI	NYREVTFFVG	LYKIFDEILV	NAADNKQRDP	KMSCIRVTID	PENNLISIWV	120
	NGKGIPVVEH	KVEKMYVPAL	IFGQLLTSSN	YDDDEKKVTG	GRNGYGAKLC	NIFSTKFTVE	180
50	TASREYKKMF	KQTWMDNMGR	AGEMELKPFN	GEDYTCTITFQ	PDLSKPKMQS	LDKDIVALMV	240
	RRAYDIAGST	KDVKVFLNGN	KLPVKGFRSY	VDMYLLKDKL	ETGNSLKVIV	EQVNRHWEVC	300
	LTMSEKGFQQ	ISFVNSIATS	KGGRHVVDYA	DQIVTKLVDV	VKKKNKGGVA	VKAHQVKNHM	360
	WIFVNALIEN	PTFDSQTKEN	MTLQPKSFGS	TCQLSEKFIK	AAIGCGIVES	ILNWVKFKAQ	420
55	VQLNKKCSAV	KHNRIKGIPK	LDDANDAGGR	NSTECTLILT	EGDSAKTLAV	SGLGVVGRDK	480
	YGVFPLRGKI	LNVRASHKQ	IMENAEINNI	IKIVGLQYKK	NYEDEDLKT	LRYGKIMIMT	540
	DQDQDGSNIK	GLLINFIIHN	WPSLLRHRFL	BEFITPIVKV	SKNKQEMAFY	SLPEFEWKS	600
	STPNHKKWKV	KYYKGLGTST	SKEAKEYFAD	MKRHRIQFKY	SGPEDDAAIS	LAFSKKQIDD	660
	RKEWLTNFM	DRRQRKLLGL	PEDYLYGQTT	TYLTYNDFIN	KELILFSNSD	NERSIPSMVD	720
60	GLKPGQRKVL	FTCFKRNDKR	EVKVAQLAGS	VAEMSSYHHG	EMSLMMTIIN	LAQNFVGSNN	780
	LNLLQPIGQF	GTRLHGGKDS	ASPRYIFTML	SSLARLLFPP	KDDHTLKFLY	DDNQRVPEW	840
	YIPIIPMVLI	NGAEGIGTGW	SCKIPNFDVR	EIVNNIRRLM	DGEEPLPMLP	SYKNFKGTIE	900
	ELAPNQYVIS	GEVAILNSTT	IBISELPVRT	WTQTYKEQVL	EPMLNGTEKT	PPLITDYREY	960
	HTDITVKFVV	KMTEEKLAEA	ERVGLHKVFK	LQTSLTCSNM	VLFDFHVGCLK	KYDITVLDILR	1020
	DFFELRLKYY	GLRKEWLLGM	LGAESAKLNE	QARFILEKID	GKII IENKPK	KELIKVLIQR	1080
65	GYSDDPVKAW	KEAQKQVPDE	EENEESSDNEK	ETEKSDSVTD	SGPTFNLYLD	MPLWYLTKEK	1140
	KDELCRLRNE	KEQELDTLKR	KSPSDLWKED	LATFIEELEA	VEAKEKQDEQ	VGLPGKGGKA	1200
	KGKKTQMAEV	LSPRGRQRFI	PRITIEKMAE	AEKKNKKKIK	NENTEGSPQE	DGVELEGLKQ	1260
	RLEKKQKREP	GTKTKKQTTL	AFKPIKKGKK	RNPWPDSESD	RSSDESNDV	PPRETEPRRA	1320
70	ATKTKFTMDL	DSDEDFSDFD	EKTDDDFVFP	SDASPPKTKT	SPKLSNKKELK	PQKSVVSDLE	1380
	ADDVKGVSPL	SSSPPATHFP	DETEITNFPV	KKNVTVKKTA	AKSQSSTSTT	GAKKRAAPKG	1440
	TKRDPALNSG	VSQKPDPAKT	KNRRKRKPST	SDDSDSNFEK	IVSKAVTSKK	SKGESDDFHM	1500
	DFDSAVAPRA	KSVRAKKPIK	YLEESDEDDL	F			

Seq ID NO: 179 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 148-7095

80	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
85	CAGCTCCTCT	GTGTTTGCCG	CCTGATTTGG	GCTAAGGATG	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420



	AACACATTCA	TTCATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTGAGCGGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
5	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
	GATTTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
10	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTCCTTGTG	AAGTTCTTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTCGAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTTCATGA	AGCAGTTTGT	1080
	AGTTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
15	GGTGCTATTTC	TCAATAATT	GCTACCCCAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
20	AACCAATCA	GGAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTCAAT	TCCACTTCCC	AACCAGTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAAGTG	TTCTTAGATC	TCCACATATG	1800
25	AACTTGTGCG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAGACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAGTAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
30	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
	TCCTTTTCTG	CAGGCCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CCTTTGCCTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACG	2400
35	GTATACAATG	GTGAGACACC	TCTTCAACCT	TCTTACAGTA	GTGAAGTCTT	TCCTCTAGTC	2460
	ACCCCTTTGT	TGCTTGACAA	TCAGATCCTC	AACACTACCC	CTGCTGCTTC	AAGTAGTGAT	2520
	TCGGCCTTGC	ATGCTACGCC	TGTATTTCCC	AGTGTGATG	TGTCATTTGA	ATCCATCCTG	2580
	TCTTCTCTATG	ATGGTGACCC	TTTGCTTCCA	TTTTCTCTCTG	CTTCTTCAG	TAGTGAATTG	2640
	TTTCGCCATC	TGCATACAGT	TTCTCAAATC	CTTCCACAAG	TTACTTCAGC	TACCGAGAGT	2700
40	GATAAGGTGC	CCTTGCATGC	TTCTCTGCCA	GTGGCTGGGG	GTGATTGCT	ATTAGAGCCC	2760
	AGCTTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGGTAGTG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACACCCC	2880
	AGCAGTGATG	CCATGATGCA	TGCACGTTCT	TCAGGGCCTG	AACCTTCTTA	TGCCTTGTCT	2940
	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTTTCTTACA	GTTCTGCAAT	ACCTGTGCAT	3000
45	GATTCTGTGG	GTGTAACCTA	TCAGGGTTCC	TTATTTAGCG	GCCCTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCAACT	GTTTCTGTAG	CTGAATTTAC	ATATACAACA	3240
	TCTGTGTTTG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGGAAATGAG	3300
50	ACTGAAGTGC	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGGCAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAT	AACCTTTTCA	TTCAACCTAC	ACATACTGTC	3540
	TCTCAAGCAT	CTGGTGACAC	TTGCTTAAAT	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
55	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATTG	CTACAACCTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAAACCTG	TCTTCCAGCT	GTGCCCAGTG	ATCCAAATAT	GGTTGAAACC	3780
	CCCAAAGTTG	ATAAAATTAG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
	AGTGAAAACA	TGCTGCACTC	TACATCTGTA	CCAGTTTTTG	ATGTGTGCGC	TACTTCTCAT	3900
60	ATGCACTCTG	CGTTCACTCA	AGGTGTGACC	ATTTCTTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTTTTGTTAA	AAAGTGAAAG	TTCCACCAAA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAAC	CAGGCCCATC	CCCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATTGATGAA	CCATTAAATA	CACATAATAA	TAAGCTTATA	4140
	CATTCCGATG	AAATTTTAAC	CTCCACCAAA	AGTTCTGTTA	CTGGTAAGGT	ATTTGCTGGT	4200
65	ATTTCCAACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTCTGTTC	TATAGGAAAT	4260
	GGGCTAGTTG	CGATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCTGTAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACCTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
	AGAGGTAGTG	ATGGCTTATC	CATTCAATAAG	TGTATGTCAT	GCTCATCCTA	TAGAGAATCA	4500
70	CAGGAAAAGG	TAATGAATGA	TTCAGACACC	CACGAAAACA	GTCTTATGGA	TCAGAATAAT	4560
	CCAATCTCAT	ACTCACTATC	TGAGAATTCT	GAAGAAGATA	ATAGAGTCAC	AAGTGTATCC	4620
	TCAGACAGTC	AAACTGGTAT	GGACAGAAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACAATGATGG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCTGACAA	GTGATGAAGA	AAGTGGATCA	4800
75	GGGCAAGGTA	CCTCAGATAG	CCTTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTTGCA	4860
	GACACTAATG	AAAAAGATGC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AATAACTCCT	4920
	GGATTCCAC	AGTCCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAAGT	GTTCCACGTT	4980
	TCAGAGGCAG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	5040
	GAATCCGAGA	AGAAGGCGAT	TATACCCCTT	TGATCCGTGT	CAGCCCTGAC	TTTTATCTGT	5100
80	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	5220
	ATTTTCAGATG	ATGTCGGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTAA	CCAGGAAGTG	5340
	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	GCAGACAGCT	CCAACCACCC	AGACAACAAG	5400
85	CACAAGAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	5520
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTT	5580
	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	5640



AAAGGAAGGA GAAATGTGA TCAGTACTGG CCTGCCGATG GGAGTGAGGA GTACGGGAAC 5700  
 TTTCTGGTCA CTCAGAAAGAG TGTGCAAGTG CTTGCCTATT ATACTGTGAG GAATTTTACT 5760  
 CTAAGAAACA CAAAAATAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACGTGTGGTC 5820  
 ACACAGTATC ACTACACGCA GTGGCCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG 5880  
 5 CTGACCTTTG TGAGAAAGGC AGCCTATGCC AAGCGCCATG CAGTGGGGCC TGTGTCTGTC 5940  
 CACTGCAGTG CTGGAGTTGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGCAG 6000  
 CAGATTCAAC ACGAAGGAAC TGTCAACATA TTTGGCTTCT TAAAAACACAT CCGTTCACAA 6060  
 AGAAATTATT TGGTACAAAC TGAGGAGCAA TATGTCTTCA TTCATGATAC ACTGGTTGAG 6120  
 10 GCCATACTTA GTAAAGAAAC TGAGGTGCTG GACAGTCATA TTCATGCCTA TGTTAATGCA 6180  
 CTCCTCATTC CTGGACCAGC AGGCAAAACA AAGCTAGAGA AACAATTCCA GCTCCTGAGC 6240  
 CAGTCAAATA TACAGCAGAG TGACTATTCT GCAGCCCTAA AGCAATGCAA CAGGGAAAAG 6300  
 AATCGAACTT CTTCTATCAT CCCTGTGGAA AGATCAAGGG TTGGCATTTC ATCCCTGAGT 6360  
 GGAGAAGGCA CAGACTACAT CAATGCCTCG TATATCATGG GCTATTACCA GAGCAATGAA 6420  
 15 TTCATCATTA CCCAGCACCC TCTCCTTCAT ACCATCAAGG ATTTCTGGAG GATGATATGG 6480  
 GACCATAATG CCCAACTGGT GGTATGATT CCTGATGGCC AAAACATGGC AGAAGATGAA 6540  
 TTTGTTTACT GGCCAAATAA AGATGAGCCT ATAAATTGTG AGAGCTTTAA GGTCACCTCT 6600  
 ATGGCTGAAG AACACAAATG TCTATCTAAT GAGGAAAAAC TTATAATTCA GGACTTTATC 6660  
 TTAGAAGCTA CACAGGATGA TTATGTACTT GAAGTGAGGC ACTTTCAGTG TCCTAAATGG 6720  
 CCAAATCCAG ATAGCCCCAT TAGTAAACTT TTTGAACCTA TAAGTGTTAT AAAAGAAGAA 6780  
 20 GTGCGCAATA GGGATTGGCC TATGATTGTT CATGATGAGC ATGGAGGAGT GACGGCAGGA 6840  
 ACTTCTGTG CTCTGACAAAC CCTTATGCAC CAACTAGAAA AAGAAAATTC CGTGGATGTT 6900  
 TACCAGGTAG CCAAGATGAT CAATCTGATG AGGCCAGGAG TCTTTGCTGA CATTGAGCAG 6960  
 TATCAGTTTC TCTACAAAGT GATCCTCAGC CTTGTGAGCA CAAGGCAGGA AGAGAATCCA 7020  
 TCCACCTCTC TGGACAGTAA TGGTGCAGCA TTGCCTGATG GAAATATAGC TGAGAGCTTA 7080  
 25 GAGTCTTTAG TTTAACACAG AAAGGGGTGG GGGGACTCAC ATCTGAGCAT TGTTCCTCCT 7140  
 TTCTTAAAT TAGGCAGGAA AATCAGCTTA GTTCTGTTAT CTGTTGATTT CCCATCACCT 7200  
 GACAGTAACT TTCATGACAT AGGATTCTGC CGCCAAATTT ATATCATTAA CAATGTGTGC 7260  
 CTTTTTGCAA GACTTGTAAT TTACTTATTA TGTTTGAAC AAAATGATTG AATTTTACAG 7320  
 TATTTCTAAG AATGGAATTG TGGTATTTTT TTCTGTATTG ATTTTAACAG AAAATTTCAA 7380  
 30 TTTATAGAGG TTAGGAATTC CAAACTACAG AAAATGTTTG TTTTATAGTG CAAATTTTTA 7440  
 GCTGTATTTG TAGCAATTAT CAGGTTTGCT AGAAATATAA CTTTAAATAC AGTAGCCTGT 7500  
 AAATAAAACA CTCTTCCATA TGATATTCAT CATTTTACAA CTGCAGTATT CACCTAAAGT 7560  
 AGAAATAATC TGTTACTTAT TGTAATAACT GCCCTAGTGT CTCCATGGAC CAAATTTATA 7620  
 35 TTTATAATTG TAGATTTTAA TATTTTACTA CTGAGTCAAG TTTTCTAGTT CTGTGTAATT 7680  
 GTTTAGTTTA ATGACGTAGT TCATTAGCTG GTCTTACTCT ACCAGTTTTC TGACATTGTA 7740  
 TTGTGTTACC TAAGTCATTA ACTTTGTTTC AGCATGTAAT TTTAACTTTT GTGGAAAATA 7800  
 GAAATACCTT CATTTTGAAA GAAGTTTTTA TGAGAATAAC ACCTTACCAA ACATTGTTCA 7860  
 AATGGTTTTT ATCCAAGGAA TTGCAAAAAT AAATATAAAT ATTGCCATTA AAAAAAATAA 7920  
 40 AAAAAAATAA AAAAAAATAA AAAAA

Seq ID NO: 180 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 45 | | | | |  
 MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKKL KFQGWDKTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFWH GKCNMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEBAV KGKGLRLALS 180  
 50 ILFEVGTENN LDFKAIIDGV ESVSFRGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVLMMDY LQNNFREQQY KFSRQVFSSY 300  
 TGKEEIEHEAV CSSEPEENVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360  
 HEFLTIDGYQD LGAILNNLLP NMSYVLQIVA ICTNGLYGY SDQLIVDMPT DNPDLDFPE 420  
 LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTKYNEAKTN 480  
 55 RSPTRGSEFS GKGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVTELEPHT VEGTSASLND 540  
 GSKTVLRSPH MNLSGTAESE NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSMEG NVWFPSSDI 660  
 TAQPDVSGSR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTPHAFPT SSRQQLVST VNVVYSQTTQ PVYNGETPLQ PSYSSEVFPL VTPLLLDNQI 780  
 60 LNTTPAASSS DSALHATPVF PSVDVSFESI LSSYDGAPLL PFSSASFSSE LFRHLHTVSQ 840  
 ILPQVTSATE SDKVPLHASL PVAGGDLLE PSLAQYSDVL STTHAASETL EFGSESGVLY 900  
 KTLMFSSQVEP PSSDAMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG 960  
 SLFSGPSHIP IPKSSLTPT ASLLQPTHAL SGDGEWSGAS SDSEFLLPDT DGLTALNISS 1020  
 PVSVAEFTYT TSFVGDDNKA LSKSEIIYGN ETELQIPSFN EMVYPSESTV MPNMYDNVKN 1080  
 65 LNASLQETSV SISSTKGMFP GSLAHTTKV FDHEISQVPE NNFSVQPTHV VSQASGDTSL 1140  
 KPVLSANSEP ASSDPASSEM LSPSTQLLFY ETSASFSTEV LLQPSFQASD VDTLLKTVLP 1200  
 AVPSDPILVE TPKVDKISST MLHLIVNSNA SSENMLHSTS VPVFDVSPTS HMHSASLQGL 1260  
 TISYASEKYE PVLLKSESSH QVVPSLYSND ELFQTANLEI NQAHPEKGRH VFATPVLSID 1320  
 EPLNLTINKL IHSDEILTST KSSVTGKVFA GIPTVASDTF VSTDHVSPIG NGHVAITAVS 1380  
 70 PHRDGVSST KLLFPSKATS ELSHSAKSDA GLVGGGEDGD TDDDDDDDDD DRGSDGLSIH 1440  
 KCMSCSSYRE SQEKVMNDSD NPISYSLSEN SEEDNRVTSV SSDSQTGMDR 1500  
 SPGKSPSANG LSQKHNDGKE ENDIQTGSAL LPLSPESKAW AVLTSDDEESG SGQGTSDSLN 1560  
 ENETSTDFSF ADTNEKDADG ILAAGDSEIT PGFPQSPTSS VTSENSEVFH VSEAEASNSS 1620  
 HESRIGLAEG LESEKKAVIP LVIVSALTFI CLVVLVGILI YWRKCFQTAH FYLEDSTSPR 1680  
 75 VISTPPTPIF PISDDVGAIP IKHFPKHVAD LHASSGFTEE FETLKEFYQE VQSCTVDLGI 1740  
 TADSSNHPDN KHKNRVINIV AYDHSRVKLA QLAEKDGKLT DYINANYVDG YNRPKAYIAA 1800  
 QGPLKSTAED FWRMIWEHNV EVIVSMITNLV EKGRKCDQY WPADGSEBYG NFLVTQKSVQ 1860  
 VLAYYTVRNF TLRNTKIKKG SQKGRPSGRV VTQYHYTQWP DMGVPEYSLP VLTFRKAAAY 1920  
 AKRHAVGPVV VHCSAGVGRT GTYIVLDSML QQIQHEGTVN IFGFLKHRS QRNYLVQTEE 1980  
 80 QYVFIHDTLV EAILSKEDEV LDSEIHAYVN ALLIPGPAGK TKLEKQFQLL SQSNIQSDY 2040  
 SAALKQCNRN KNRSTSSIPV ERSRVGISSL SGEPTYDINA SYIMGYYSN EFTITQHPLL 2100  
 HTIKDFWRMI WDHNAQLVVM IPDGQNMAD EFWYWPKNDE PINCESFKVT LMAEEHKCLS 2160  
 NEEKLIIQDF ILEATQDDYV LEVRHFQCPK WPNPDSPIK TFELESVIKE EAANRDGPMI 2220  
 VHDEHGGVTA GTFECALTLM HQLEKENSVD VYQVAKMINL MRPGVFADIE QYQFLYKVIL 2280  
 85 SLVSTRQEEN PSTSLDSNGA ALPDGNIAES LESLV

Seq ID NO: 181 DNA sequence  
 Nucleic Acid Accession #: Eos sequence



	1	11	21	31	41	51	
5	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
10	CTTGTTGAAG	AGATTGGCTG	GTCTATACAA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
	AACACATTCA	TTCATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTGAGCGGAG	GAGTTTCAGA	AATGGTGTTT	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
15	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
	GATTTTCAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
20	AATGGCTCAT	TGCATGCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTTTGTG	AAGTTCTTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTCGAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCATGA	AGCAGTTTGT	1080
	AGTTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
25	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGACGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTG	TCAATAAATT	GCTACCCAAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGTG	AAGAAATAAT	CAAGGAGGAG	1440
30	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCAGTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AAC TGCCACC	TCACACTGTG	1740
35	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAAGTG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTCGG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCCTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCCTCC	1980
	GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
40	GAAGATTCAA	CTTCATCAGG	TTTCAAGAGG	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGGATG	AATCTGAGAA	GACAACCAAG	2220
	TCCTTTTCTG	CAGGCCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CCTTTGCCCTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
45	TCCAGACAAC	AGGATTGGT	CTCCACGGTG	AACGTGGTAT	ACTCGCAGAC	AACCAACCG	2400
	GTATACAAATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAAGGC	AGTTATACCC	CTTGTGATCG	TGTCAGCCCT	GACTTTTATC	2520
	TGTCTAGTGG	TTCTTGTGGG	TATTCTCATC	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
	TTTTACTTAG	AGGACAGTAC	ATCCCTTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTT	2640
50	CCAATTTTCAG	ATGATGTCGG	AGCAATTCCA	ATAAAGCACT	TTCCAAAGCA	TGTTGCAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTAGAAGA	TTTGAGACAC	TGAAAGAGTT	TTACCAGGAA	2760
	GTGCAGAGCT	GTACTGTTGA	CTTAGGTATT	ACAGCAGACA	GCTCCAACCA	CCCAGACAAC	2820
	AAGCACAAGA	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAAGCTAGCA	2880
	CAGCTTGCTG	AAAAGGATGG	CAAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
55	TACAACAGAC	CAAAAGCTTA	TATTGCTGCC	CAAGGCCAC	TGAAATCCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAGTTATTG	TCATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAAATG	TGATCAGTAC	TGGCCTGCCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACTTTCTGG	TCACTCAGAA	GAGTGTGCAA	GTGCTTGCC	ATTATACTGT	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAAAAT	AAAAAAGGGC	TCCAGAAAG	GAAGACCCAG	TGGACGTGTG	3240
60	GTCACACAGT	ATCACTACAC	GCAGTGGCCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGTGAGAAA	GGCAGCCTAT	GCCAAGCGCC	ATGCAGTGGG	GCCTGTTGTC	3360
	GTCCACTGCA	GTGCTGGAGT	TGGAAGAACA	GGCACATATA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTC	AACACGAAGG	AACTGTCAAC	ATATTTGGCT	TCTTAAAACA	CATCCGTTCA	3480
	CAAGAGAAAT	ATTTGGTACA	AACTGAGGAG	CAATATGTCT	TCATTATGTA	TACACTGGTT	3540
65	GAGGCCATAC	TTAGTAAAGA	AACTGAGGTG	CTGGACAGTC	ATATTATATG	CTATGTTAAT	3600
	GCACTCCTCA	TTCCTGGACC	AGCAGGCAAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGACTAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAATCGAA	CTTCTTCTAT	CATCCCTGTG	GAAAGATCAA	GGGTGCGCAT	TTTATCCCTG	3780
	AGTGGAGAAG	GCACAGACTA	CATCAATGCC	TCCTATATCA	TGGGCTATTA	CCAGAGCAAT	3840
70	GAATTCATCA	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTCTCT	GAGGATGATA	3900
	TGGGACCATA	ATGCCCAACT	GGTGGTTATG	ATTCCTGATG	GCCAAAACAT	GGCAGAAGAT	3960
	GAATTTGTTT	ACTGGCCAAA	TAAAGATGAG	CCTATAAATT	GTGAGAGCTT	TAAGGTCACT	4020
	CTTATGGCTG	AAGAACACAA	ATGTCTATCT	AATGAGGAAA	AACTTATAAT	TCAGGACTTT	4080
	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTTGAAGTGA	GGCACTTTCA	GTGTCCTAAA	4140
75	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAAGAA	4200
	GAAGCTGCCA	ATAGGGATGG	GCCTATGATG	GTTCAATGAT	AGCATGGAGG	AGTGACGGCA	4260
	GGAACTTTCT	GTGCTCTGAT	AACCTTATTG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCAGG	TAGCCAAGAT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATTGAG	4380
	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTGA	GCACAAGGCA	GGAAGAGAAT	4440
80	CCATCCACCT	CTCTGGACAG	TAATGGTGCA	GCATTGCCTG	ATGGAAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAAC	CAGAAAGGGG	TGGGGGGACT	CACATCTGAG	CATTGTTTTT	4560
	CTCTTCCTAA	AATTAGGCAG	GAAAATCAGT	CTAGTTCTGT	TATCTGTTGA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTTCATGA	CATAGGATTC	TGCCGCCAAA	TTTATATCAT	TAACAATGTG	4680
	TGCCTTTTTG	CAAGACTTGT	AATTTACTTA	TTATGTTTGA	ACTAAAATGA	TTGAATTTTA	4740
85	CAGTATTTCT	AAGAATGGAA	TTGTGGTATT	TTTTTCTGTA	TTGATTTTAA	CAGAAAATTT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAAACTA	CAGAAAATGT	TTGTTTTTAG	TGTCAAATTT	4860
	TTAGCTGTAT	TTGTAGCAAT	TATCAGGTTT	GCTAGAAATA	TAACTTTAA	TACAGTAGCC	4920
	TGTAAATAAA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCAGT	ATTACCTTAA	4980



AGTAGAATA ATCTGTTACT TATTGTAAAT ACTGCCCTAG TGTCTCCATG GACCAAATTT 5040  
 ATATTTTATAA TTGTAGATT TTATATTTTA CTACTGAGTC AAGTTTTCTA GTTCTGTGTA 5100  
 ATTGTTTAGT TTAATGACGT AGTTCATTAG CTGGTCTTAC TCTACCAGTT TTCTGACATT 5160  
 GTATTGTGTT ACCTAAGTCA TTAACCTTGT TTCAGCATGT AATTTTAACT TTTGTGGAAA 5220  
 ATAGAAATAC CTTCAATTTG AAAGAAGTTT TTATGAGAAT AACACCTTAC CAAACATTGT 5280  
 TCAAATGGTT TTTATCCAAG GAATTGCAAA AATAAATATA AATATTGCCA TTAACAAAAA 5340  
 AAAAAAAAAA AAAAAAAAAA AAAAAA

Seq ID NO: 182 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | |  
 15 MRILKRFLAC IQLLCVCRLD WANGYYRQOR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKKL KFGQWDKTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFWH GKNMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEEAV KGKGKLRALS 180  
 ILFEVGTEN LDFKAIIDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 20 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVMLMDY LQNNFREQQY KFSRQVFSSY 300  
 TGKEEIEHAV CSSEPENVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360  
 HEFLTGDYQD LGAILNNLLP NMSYVLQIVA ICTNGLYGKY SDQLIVDMPT DNPELDFPE 420  
 LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDPVNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPTH VEGTSASLND 540  
 25 GSKTVLRSPH MNLSTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSMEG NVWFPSSTDI 660  
 TAQPDVSGSR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTPHAFTP SSRQQLVST VNVVYSQTTQ PVYNAEASNS SHESRIGLAE GLESEKKAVI 780  
 PLVIVSALTF ICLVVLVGIL IYWRKCFQTA HFYLEDSTSP RVISTPPTPI FPISTDDVGA 840  
 30 PIKHFPHKVA DLHASSGFTE EFETLKEFYQ EVQSVTVDLG ITADSSNHPD NKHKNRYINI 900  
 VAYDHSRVKL AQLAEKDGKL TDYINANYVD GYNRPKAYIA AQGPLKSTAE DFWRMIWEHN 960  
 VEVIVMITNL VEKGRKCDQ YWPADGSEY GNFLVTQKSV QVLAYYTVRN FTLRNTKIKK 1020  
 GSQKGRPSGR VVTQYHYTQW PDMGVPEYSL PVLTFVRKAA YAKRHAVGPV VVHCSAGVGR 1080  
 TGTIYVLDSM LQIQHEGTV NIFGFLKHIR SQRYNLVQTE EQYVFIHDTL VEAILSKETE 1140  
 35 VLDSHIHAYV NALLIPGPAG KTKLEKQFQL LSQSNIQQSD YSAALKQCNR EKNRTSSIIP 1200  
 VERSRVGISS LSGEGTDYIN ASYIMGYYSQ NEFIITQHPL LHTIKDFWRM IWDHNAQLVV 1260  
 MIPDGQNAE DEFVYWPND EPINCESFKV TLMAEHKCL SNEEKLIQD FILEATQDDY 1320  
 VLEVRHFQCP KWPNDSPIS KTFELISVIK EEAANRDGPM IVHDEHGGVT AGTFCALTTL 1380  
 MHQLEKENS DVYQVAKMIN LMRPGVFADI EQYQFLYKVI LSLVSTRQEE NPSTSLDSNG 1440  
 AALPDGNIAE SLESIV

Seq ID NO: 183 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 148-4494

1 11 21 31 41 51  
 | | | | |  
 45 CACACATACG CACGCACGAT CTCACTTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAA AAC ATTTCTCTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCCT CGCTTGCAAT 180  
 50 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTTGTTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAA TTGGGGAAAG 300  
 AAATATCCAA CATGTAATAG CCCAAAAACA TCTCCTATCA ATATTGATGA AGATCTTACA 360  
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTGAGGGTT GGGATAAAAC ATCATTGGAA 420  
 55 AACACATTCA TTCATAACAC TGGGAAAAACA GTGGAAATTA ATCTCACTAA TGAATACCGT 480  
 GTCAGCGGAG GAGTTTCAGA AATGGTGTGT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 AAATGCAATA TGTCTCTGTA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600  
 GAGATGCAAA TCTACTGCTT TGATGCAGAG CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660  
 GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAATTTG 720  
 60 GATTTCAAAG CGATTATTGA TGGAGTCGAA ATGTTTAGTC GTTTTGGGAA GCAGGCTGCT 780  
 TTAGATCCAT TCATACTGTT GAACCTTCTG CCAAACTCAA CTGACAAAGTA TTACATTTAC 840  
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTTTGTG AAGTCTTAC AATGCAACAA 960  
 TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTCCAGAGCA ACAGTACAAG 1020  
 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTCATGA AGCAGTTTGT 1080  
 65 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA 1140  
 TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGACAG TTTGTACCAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATTG TCAATAAATT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 70 TGCATAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGCAGAT GCCTACTGAT 1380  
 AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAAC TGAGAAATAA CAAGGAGGAG 1440  
 GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500  
 AACCAATCA GGAAAAAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560  
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620  
 75 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCAGTCAC TAAATTAGCC 1680  
 ACAGAAAAAG ATATTTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAAC TGCTTAGATC TCCACATATG 1800  
 AACTTGTCGG GGAATGACAG ATCCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860  
 AGTTTATTGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATCTTTCAGG CTCCAGTCCC 1920  
 80 GCAACTTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980  
 GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040  
 GAAGATTCAA CTTTATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAAT 2100  
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTTCTCC AGACTAATA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220  
 85 TCCTTTCTCG CAGGCCAGT GATGTACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA 2280  
 CATTATTCTA CCTTGCCTA CTTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340  
 TCCAGACAAC AGGATTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG 2400  
 GTATACAATG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTTG 2460



	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
5	ATTTTCAGATG	ATGTCGGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGGAAGTGC	AGAGCTGTAC	TGTTGACTTA	2760
	GGTATTACAG	CAGACAGCTC	CAACCACCCA	GACAACAAGC	ACAAGAATCG	ATACATAAAT	2820
	ATCGTTGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCAAA	2880
	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
10	GCTGCCCCAAG	GCCCCACTGAA	ATCCACAGCT	GAAGATTTCCT	GGAGAATGAT	ATGGGAACAT	3000
	AATGTGGAAG	TTATTGTTCAT	GATAACAAAC	CTCGTGGAGA	AAGGAAGGAG	AAAATGTGAT	3060
	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAACT	TTCTGGTCAC	TCAGAAGAGT	3120
	GTGCAAGTGC	TTGCCTATTA	TACTGTGAGG	AATTTTACTC	TAAGAAACAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGA	CGTGTGGTCA	CACAGTATCA	CTACACGCAG	3240
	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGGCA	3300
15	GCCTATGCCA	AGCGCCATGC	AGTGGGGCCT	GTTGTCTGTC	ACTGCAGTGC	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAAC	3420
	GTCAACATAT	TTGGCTTCTT	AAAACACATC	CGTTCACAAA	GAAATTATTT	GGTACAAACT	3480
	GAGGAGCAAT	ATGTCTTCAT	TCATGATACA	CTGGTTGAGG	CCATACTTAG	TAAAGAAACT	3540
20	GAGGTGCTGG	ACAGTCAATAT	TCATGCCTAT	GTTAATGCAC	TCCTCATTCC	TGGACCAGCA	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCCAG	CTCCTGAGCC	AGTCAAAATAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGAAAAAGA	ATCGAACTTC	TTCTATCATC	3720
	CCTGTGGAAA	GATCAAGGGT	TGGCATTTCA	TCCCTGAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGCCTCCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCATTAC	CCAGCACCTT	3840
25	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	ACCATAATGC	CCAAGTGGTG	3900
	GTTATGATTG	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACTCTTA	TGGCTGAAGA	ACACAAATGT	4020
	CTATCTAATG	AGGAAAAACT	TATAATTACG	GACTTTATCT	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	CAAATCCAGA	TAGCCCCATT	4140
30	AGTAAAACTT	TTGAACCTAT	AAGTGTATA	AAAGAAGAAG	CTGCCAATAG	GGATGGGCCT	4200
	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	CTTCTGTGTC	TCTGACAACC	4260
	CTTATGCACC	AACTAGAAAA	AGAAAAATCC	GTGGATGTTT	ACCAGGTAGC	CAAGATGATC	4320
	AATCTGATGA	GGCCAGGAGT	CTTGTCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACCTCTCT	GGACAGTAAT	4440
35	GGTCAGCAT	TGCCCTGATG	AAATATAGTG	GAGAGCTTAG	AGTCTTTAGT	TTAACACAGA	4500
	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTTCTCTT	TCCTAAAATT	AGGCAGGAAA	4560
	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTC	CCATCACCTG	ACAGTAACTT	TCATGACATA	4620
	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	TTTTTGCAAG	ACTTGTAAAT	4680
	TACTTATTAT	GTTTGAACCTA	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
40	GGTATTTTTT	TCTGTATTGA	TTTAAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTCC	4800
	AAACTACAGA	AAATGTTTGT	TTTATGTGTC	AAATTTTAG	CTGTATTTGT	AGCAATTATC	4860
	AGGTTTGCTA	GAAATATAAC	TTTAAATACA	GTAGCCTGTA	AATAAAACAC	TCTTCCATAT	4920
	GATATTCAAC	ATTTTACAAC	TGCAGTATTC	ACCTAAAGTA	GAAATAATCT	GTTACTTATT	4980
	GTAAATACTG	CCCTAGTGTG	TCCATGGACC	AAATTTATAT	TTATAATTGT	AGATTTTTAT	5040
45	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATTG	TTTAGTTTAA	TGACGTAGTT	5100
	CATTAGCTGG	TCTTACTCTA	CCAGTTTCTT	GACATTGTAT	TGTGTTACCT	AAGTCATTAA	5160
	CTTTGTTTCA	GCATGTAATT	TTAACTTTTG	TGGAAAATAG	AAATACCTTC	ATTTTGAAAG	5220
	AAGTTTTTAT	GAGAATAACA	CCTTACCATA	CATTGTTCAA	ATGGTTTTTA	TCCAAGGAAT	5280
	TGCAAAAATA	AATATAAATA	TTGCCATTAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	5340
50	AAA						

Seq ID NO: 184 Protein sequence:  
Protein Accession #: EOS sequence

	1	11	21	31	41	51	
55							
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLKKL	KFQGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
	FKASKITFWH	GKCNMSSDGS	EHSLEGQKFP	LEMQUIYCFDA	DRFSSFEEAV	KGKGKLRALS	180
60	ILFEVGTENN	LDFKALIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVFCEVLTMQ	QSGYVMLMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGKEEIHAEV	CSSEPEENVQA	DPENYTSLLV	TWERPRVVYD	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNNLLP	NMSYVLQIVA	ICTNGLYGKY	SDQLIVDMPT	DNPELDLFPE	420
	LIGTEEIIKE	EEEGKDIEEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTTHYNRI	GTYKNEAKTN	480
65	RSPTRGSEFS	GKGDVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPhT	VEGTSASLND	540
	GSKTVLRSPH	MNLSGTAESL	NTVSITEYEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETITYD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NVWFPSSTDI	660
	TAQPDVSGSR	ESFLQTNYTE	IRVDESEKTT	KSFSAGPVMS	QGPSVTDLEM	PHYSTFAYFP	720
	TEVTPHAFTP	SSRQQDLVST	VNVVYSQTTQ	PVYNEASNSS	HESRIGLAEG	LESEKKAVIP	780
70	LVIVSALTFI	CLVVLVGILI	YWRKCFQTAH	FYLEDSTSPR	VISTPPTPIF	PISDDVGAIP	840
	IKHFPKHVAD	LHASSGFTEE	FEEVQSCTVD	LGITADSSNH	PDNKHKNRYI	NIVAYDHSRV	900
	KLAQLAEKDG	KLTDYINANY	VDGYNRPKAY	IAAQGPLKST	AEDFWRMWE	HNVEVIVMIT	960
	NLVEKGRRKC	DQYWPADGSE	EYGNFLVTQK	SVQVLAYYTV	RNFTRLNRTKI	KKGSQKGRPS	1020
	GRVVTQYHYT	QWPDMGVPEY	SLPVLTFVRK	AAAYAKRHAVG	PVVVHCSAGV	GRTGTYIVLD	1080
75	SMLQQIQHEG	TVNIFGFLKH	IRSQRNYLVQ	TEEQYVFIHD	TLVEAILSKE	TEVLDSHIHA	1140
	YVNALLIPGP	AGKTKLEKQF	QLLSQSNIQQ	SDYSAALKQC	NREKNRTSSI	IPVERSRVGI	1200
	SSLSGEGTDY	INASYIMGY	QSNEFIITQH	PLLHTIKDFW	RMIWDHNAQL	VVMIPDGQNM	1260
	AEDFVYWPVN	KDEPINCESF	KVTLMABEHH	KSNNEEKLI	QDFILEATQD	DYVLEVRHFQ	1320
	CPKWPNDPSP	ISKTFELISV	IKEEAANRDG	PMIVHDEHGG	VTAGTFCALT	TLMHQLEKEN	1380
80	SVDVYQVAKM	INLMRPGVFA	DIEQYQFLYK	VILSLVSTRO	EENPSTSLDS	NGAALPDGNI	1440
	AESLESLEV						

Seq ID NO: 185 DNA sequence  
Nucleic Acid Accession #: EOS sequence  
Coding sequence: 501-4514

1 11 21 31 41 51



	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCCTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCATT	180
5	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAT	TGGGGAAAGA	300
	AATATCCAAC	ATGTAATAGC	CCAAAACAAT	CTCCTATCAA	TATTGATGAA	GATCTTACAC	360
	AAGTAAATGT	GAATCTTAAG	AACTTTAAAT	TTGAGGGTTG	GGATAAAACA	TCATTGGAAA	420
	ACACATTCAT	TCATAACACT	GGGAAAACAG	TGGAAATTAA	TCTCACTAAT	GACTACCGTG	480
10	TCAGCGGAGG	AGTTTCAGAA	ATGGTGTTTA	AAGCAAGCAA	GATAACTTTT	CACTGGGGAA	540
	AATGCAATAT	GTCATCTGAT	GGATCAGAGC	ATAGTTTAGA	AGGACAAAAA	TTTCCACTTG	600
	AGATGCAAAT	CTACTGCTTT	GATGCGGACC	GATTTTCAAG	TTTTGAGGAA	GCAGTCAAAG	660
	GAAAAGGGAA	GTTAAGAGCT	TTATCCATTT	TGTTTGAGGT	TGGGACAGAA	GAAAATTTGG	720
	ATTTCAAAGC	GATTATTGAT	GGAGTCGAAA	GTGTTAGTCG	TTTTGGGAAG	CAGGCTGCTT	780
15	TAGATCCATT	CATACTGTTG	AACCTTCTGC	CAAACTCAAC	TGACAAGTAT	TACATTTACA	840
	ATGGCTCATT	GACATCTCCT	CCCTGCACAG	ACACAGTTGA	CTGGATTGTT	TTTAAAGATA	900
	CAGTTAGCAT	CTCTGAAAGC	CAGTTGGCTG	TTTTTTGTGA	AGTTCTTACA	ATGCAACAAT	960
	CTGGTTATGT	CATGCTGATG	GACTACTTAC	AAAACAATTT	TCGAGAGCAA	CAGTACAAAGT	1020
	TCTCTAGACA	GGTGTTCCTC	TCATACACTG	GAAAGGAAGA	GATTTCATGAA	GCAGTTTGTA	1080
20	GTTCAGAAC	AGAAAATGTT	CAGGCTGACC	CAGAGAATTA	TACCAGCCTT	CTTGTTACAT	1140
	GGGAAAGACC	TCGAGTCGTT	TATGATACCA	TGATTGAGAA	GTTTGCAGTT	TTGTACCAGC	1200
	AGTTGGATGG	AGAGGACCAA	ACCAAGCATG	AAATTTTGAC	AGATGGCTAT	CAAGACTTGG	1260
	GTGCTATTCT	CAATAATTTG	CTACCCAATA	TGAGTTATGT	TCTTCAGATA	GTAGCCATAT	1320
	GCACTAATGG	CTTATATGGA	AAATACAGCG	ACCAACTGAT	TGTCGACATG	CCTACTGATA	1380
25	ATCCTGAACT	TGATCTTTTC	CCTGAATTAA	TTGGAAGTGA	AGAAATAATC	AAGGAGGAGG	1440
	AAGAGGGAAA	AGACATTGAA	GAAGGCGCTA	TTGTGAATCC	TGGTAGAGAC	AGTGTACAA	1500
	ACCAAATCAG	GAAAAAGGAA	CCCCAGATTT	CTACCACAAC	ACACTACAAT	CGCATAGGGA	1560
	CGAAATACAA	TGAAGCCAAG	ACTAACCGAT	CCCCAACAAAG	AGGAAGTGAA	TTCTCTGGAA	1620
	AGGGTGATGT	TCCCAATACA	TCTTTAAATT	CCACTTCCCA	ACCAGTCACT	AAATTAGCCA	1680
30	CAGAAAAAGA	TATTTCTTGG	ACTTCTCAGA	CTGTGACTGA	ACTGCCACCT	CACACTGTGG	1740
	AAGGTACTTC	AGCCTCTTTA	AATGATGGCT	CTAAAACGTG	TCTTAGATCT	CCACATATGA	1800
	ACTTGTGCGG	GACTGCAGAA	TCCTTAAATA	CAGTTTCTAT	AACAGAATAT	GAGGAGGAGA	1860
	GTTTATTGAC	CAGTTTCAAG	CTTGATACTG	GAGCTGAAGA	TTCTTCAGGC	TCCAGTCCCG	1920
	CAACTTCTGC	TATCCCATTC	ATCTCTGAGA	ACATATCCCA	AGGGTATATA	TTTTCTCCCG	1980
35	AAAACCCAGA	GACAATAACA	TATGATGTCC	TTATACCAGA	ATCTGCTAGA	AATGCTTCCG	2040
	AAGATTCAAC	TTCATCAGGT	TCAGAAGAAT	CACTAAAGGA	TCCTTCTATG	GAGGGAAATG	2100
	TGTGGTTTCC	TAGCTCTACA	GACATAACAG	CACAGCCCGA	TGTTGGATCA	GGCAGAGAGA	2160
	GCTTTCTCCA	GACTAATTAC	ACTGAGATAC	GTGTTGATGA	ATCTGAGAAG	ACAACCAAGT	2220
	CCTTTTCTGC	AGGCCCACTG	ATGTCACAGG	GTCCCTCAGT	TACAGATCTG	GAAATGCCAC	2280
40	ATTATTCTAC	CTTTGCCCTAC	TTCCCAACTG	AGGTAACACC	TCATGCTTTT	ACCCATCCT	2340
	CCAGACAACA	GGATTGGTGC	TCCCAAGTCA	ACGTGGTATA	CTCGCAGACA	ACCCAACCGG	2400
	TATACAATGA	GGCCAGTAAT	AGTAGCCATG	AGTCTCGTAT	TGGTCTAGCT	GAGGGGTTGG	2460
	AATCCGAGAA	GAAGGCAGTT	ATACCCCTTG	TGATCGTGTC	AGCCCTGACT	TTTATCTGTC	2520
	TAGTGGTTCT	TGTGGGTATT	CTCATCTACT	GGAGGAAATG	CTTCCAGACT	GCACACTTTT	2580
45	ACTTAGAGGA	CAGTACATCC	CCTAGAGTTA	TATCCACACC	TCCAACACCT	ATCTTTCCAA	2640
	TTTCAGATGA	TGTCGGAGCA	ATTCCAATAA	AGCACTTTCC	AAAGCATGTT	GCAGATTTAC	2700
	ATGCAAGTAG	TGGGTTTACT	GAAGAATTG	AGACACTGAA	AGAGTTTAC	CAGGAAGTGC	2760
	AGAGCTGTAC	TGTTGACTTA	GGTATTACAG	CAGACAGCTC	CAACCACCCA	GACAACAAGC	2820
	ACAAGAATCG	ATACATAAAT	ATCGTTGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	2880
50	TTGCTGAAAA	GGATGGCAAA	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	2940
	ACAGACCAAA	AGCTTATATT	GCTGCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTTCT	3000
	GGAGAATGAT	ATGGGAACAT	AATGTGGAAG	TTATTGTCTAT	GATAACAAAC	CTCGTGGAGA	3060
	AAGGAAGGAG	AAAATGTGAT	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAGT	3120
	TTCTGGTCAC	TCAGAAGAGT	GTGCAAGTGC	TTGCCTATTA	TACTGTGAGG	AATTTTACTC	3180
55	TAAGAAACAC	AAAAATAAAA	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGG	CGTGTGGTCA	3240
	CACAGTATCA	CTACACGCAG	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	3300
	TGACCTTTGT	GAGAAAGGCA	GCCTATGCCA	AGCGCCATGC	AGTGGGGCCT	GTTGTGCTCC	3360
	ACTGCAGTGC	TGGAGTTGGA	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	3420
	AGATTCAACA	CGAAGGAACT	GTCAACATAT	TTGGCTTCTT	AAAACACATC	CGTTCAAAA	3480
60	GAAATTATTT	GGTACAAACT	GAGGAGCAAT	ATGTCTTCAT	TCATGATACA	CTGGTTGAGG	3540
	CCATACTTAG	TAAAGAACT	GAGGTGCTGG	ACAGTCATAT	TCATGCCTAT	GTTAATGCAC	3600
	TCCTCATTTCC	TGGACCAGCA	GGCAAAACAA	AGCTAGAGAA	ACAATTCCAG	CTCCTGAGCC	3660
	AGTCAAATAT	ACAGCAGAGT	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAAGA	3720
	ATCGAACTTC	TTCTATCATC	CCTGTGGAAA	GATCAAGGGT	TGGCATTTC	TCCCTGAGTG	3780
65	GAGAAGGCAC	AGACTACATC	AATGCCCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	3840
	TCATCATTAC	CCAGCACCTC	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	3900
	ACCATATATG	CCAAGTGGTG	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	3960
	TTGTTTACTG	GCCAAATAAA	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACCTTAA	4020
	TGGCTGAAGA	ACACAAATGT	CTATCTAATG	AGGAAAAACT	TATAATTTCAG	GACTTTATCT	4080
70	TAGAAGCTAC	ACAGGATGAT	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	4140
	CAAATCCAGA	TAGCCCCATT	AGTAAACTTT	TTGAACTTAT	AAGTGTTATA	AAAGAAGAAG	4200
	CTGCCAATAG	GGATGGGCCT	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	4260
	CTTTCTGTGC	TCTGACAACC	CTTATGCACC	AACTAGAAAA	AGAAAATTCC	GTGGATGTTT	4320
	ACCAGGTAGC	CAAGATGATC	AATCTGATGA	GGCCAGGAGT	CTTGTCTGAC	ATTGAGCAGT	4380
75	ATCAGTTTCT	CTACAAAGTG	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	4440
	CCACCTCTCT	GGACAGTAAT	GGTGCAGCAT	TGCCTGATGG	AAATATAGCT	GAGAGCTTAG	4500
	AGTCTTTAGT	TTAACACAGA	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTTCTCT	4560
	TCCTAAATTT	AGCAGGAAA	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTC	CCATCACCTG	4620
	ACAGTAACTT	TCATGACATA	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	4680
80	TTTTTGCAAG	ACTTGTAATT	TACTTATTAT	GTTTGAAC	AAATGATTGA	ATTTTACAGT	4740
	ATTTCTAAGA	ATGGAATTGT	GGTATTTTTT	TCTGTATTGA	TTTAAACAGA	AAATTTCAAT	4800
	TTATAGAGGT	TAGGAATTCC	AAACTACAGA	AAATGTTTGT	TTTTAGTGTC	AAATTTTTAG	4860
	CTGTATTTGT	AGCAATTATC	AGGTTTGCTA	GAAATATAAC	TTTAAATACA	GTAGCCTGTA	4920
	AATAAAACAC	TCTTCCATAT	GATATTCAAC	ATTTTACAAC	TGCAGTATTC	ACCTAAAGTA	4980
85	GAAATAATCT	GTTACTTATT	GTAATAATCT	CCCTAGTGTC	TCCATGGACC	AAATTTATAT	5040
	TTATAATTGT	AGATTTTAT	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATTG	5100
	TTTAGTTTAA	TGACGTAGTT	CATTAGCTGG	TCTTACTCTA	CCAGTTTTCT	GACATTGTAT	5160



TGTGTTACCT AAGTCATTAA CTTTGTTCCTA GCATGTAATT TTAACCTTTTG TGGAAAATAG 5220  
 AAATACCTTC ATTTTGAAG AAGTTTTTAT GAGAATAACA CCTTACCAA CATTTGTTCAA 5280  
 ATGGTTTTTA TCCAAGGAAT TGCAAAAATA AATATAAATA TTGCCATTAA AAAAAAAAAA 5340  
 AAAAAAAAAA AAAAAAAAAA AAA

Seq ID NO: 186 Protein sequence:  
 Protein Accession #: EOS sequence

10 1 11 21 31 41 51  
 | | | | |  
 MVFKASKITF HWGKCNMSSD GSEHSLEGQK FPLEMQIYCF DADRFSSFEE AVKKGKGLRA 60  
 LSILFEVGT E NLDFAKIID GVESVSRFGK QAALDPFILL NLLPNSTDKY YIYNGSLTSP 120  
 PCTDITVDWIV FKDTVSISES QLAVFCEVLT MQQSGYVMLM DYLNQNFREQ QYKFSRQVFS 180  
 SYTGKEEIIHE AVCSSEPENV QADPENYTS LVTWERPRVV YDTMIEKFAV LYQQLDGEDQ 240  
 15 TKHEFLTDGY QDLGAILNNL LPNMSYVLQI VAICTNGLYG KYSDQLIVDM PTDNPELDLF 300  
 PELIGTEBII KEEEEEKIDIE EGAIVNPGRD SATNQIRKKE PQISTTTTHYN RIGTKYNEAK 360  
 TNRSPTRGSE FSGKGDVPNT SLNSTSQPVT KLATEKDISL TSQTVTELPP HTVEGTSASL 420  
 NDGSKTVLRS PHMNLSTAE SLNTVSITEY EEESLLTSFK LDTGAEDSSG SSPATSAIPF 480  
 ISENISQGYI FSSSENPETIT YDVLIPESAR NASEDSTSSG SEESLKDPSM EGNVWFPSST 540  
 20 DITAQPDVGS GRESFLQTNV TEIRVDESEK TTKSFSAGPV MSQGPSVTDL EMPHYSTFAY 600  
 FPTEVTPHAF TPSSRQDLV STVNVSYSQT TQPVYNEASN SSSESRLGLA EGLESEKKAV 660  
 IPLVIVSALT FICLVVLVGI LIYWRKCFQT AHFYLEDSTS PRVISTPPTP IFPISDDVGA 720  
 IPIKHFPHKV ADLHASGFT EEFETLKEFY QEVQSTVDL GITADSSNHP DNKHKNNRYIN 780  
 IVAYDHSRVK LAQLAEKDGK LTDYINANYV DGYNRPKAYI AAQGPLKSTA EDFWRMIWEH 840  
 25 NVEVIVMITN LVEKGRRKCD QYWPADGSEE YGNFLVTQKS VQVLAYYTVR NFTLRNTKIK 900  
 KGSQKGRPSG RVVTQYHYTQ WPDMDGVPEYS LPVLTFFVRKA AYAKRHAVGP VVHCSAGVG 960  
 RTGTIYVLDL MLQQIQHEGT VNIFGFLKHI RSQRNYLVQT EEQYVFIHDT LVEAILSKET 1020  
 EVLDSEIHAY VNALLIPGPA GKTKEKQFQ LLSQSNIIQS DYSALKQCN REKNRTSSII 1080  
 PVERSRVGIS SLSGEGTDYI NASYIMGYVQ SNEFIITQHP LLHTIKDFWR MIWDHNAQLV 1140  
 30 VMIPDQNMMA EDEFVYWPKN DEPINCESFK VTLMAEEHKC LSNEEKLIQ DFILEATQDD 1200  
 YVLEVRHFQC PKWPNPDSPI SKTFELISVI KEEAANRDGP MIVHDEHGGV TAGTFCALTT 1260  
 LMHQLKENS VDVYQVAKMI NLMRPGVFAD IEQYQFLYKV ILSLVSTRQE ENPSTSLDSN 1320  
 GAALPDGNIA ESLESV

Seq ID NO: 187 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 148-4632

40 1 11 21 31 41 51  
 | | | | |  
 CACACATACG CACGCACGAT CTCACTTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAAC ATTTCTCTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AACGTTTCCT CGCTTGCAAT 180  
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 45 CTTGTTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG 300  
 AAATATCCAA CATGTAATAG CCCAAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA 360  
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA 420  
 AACACATTC A TTCATAACAC TGGGAAAACA GTGGAAATTA ATCTCACTAA TGACTACCGT 480  
 GTCAGCGGAG GAGTTTCAGA AATGGTGTGT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 50 AAATGCAATA TGTCATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAA ATTTCCTACT 600  
 GAGATGCAAA TGTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660  
 GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAATTTG 720  
 GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTC GTTTTGGGAA GCAGGCTGCT 780  
 TTAGATCCAT TCATACTGTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTTAC 840  
 55 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTTTGTG AAGTCTTAC AATGCAACAA 960  
 TCTGTTATG TCTAGCTGAT GGACTACTTA CAAAACAATT TTCGAGAGCA ACAGTACAAG 1020  
 TTCTCTAGAC AGGTGTTTTT CTCATACACT GGAAAGGAAG AGATTCATGA AGCAGTTTGT 1080  
 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA 1140  
 60 TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCACT TTTGTACCAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 TGCACATAAT GCTTATATGG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380  
 AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAAGT AAGAAATAAT CAAGGAGGAG 1440  
 65 GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500  
 AACCAATCA GGAAAAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560  
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620  
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCAAGTCA TAAATTAGCC 1680  
 ACAGAAAAAG ATATTTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 70 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACCTG TTCTTAGATC TCCACATATG 1800  
 AACTTGTCTG GACTGCAGA ATCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860  
 AGTTTATTGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920  
 GCAACTTCTG CTATCCCAT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTCTCTCC 1980  
 GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040  
 75 GAAGATTCAA CTTATCAGG TTCAGAAAGG TCACATAAGG ATCTTCTAT GGAGGGAAT 2100  
 GTGTGTTTCT CTAGCTCTAC AGCAGTACAC GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTCTCTC AGACTAATTA CACTGAGATA CGTGTGTGAT AATCTGAGAA GACAACCAAG 2220  
 TCCTTTTCTG CAGGCCAGT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA 2280  
 CATTATTCTA CCTTTGCCA CTTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340  
 80 TCCAGACAAC AGGATTTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG 2400  
 GTATACAATG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTTG 2460  
 GAATCCGAGA AGAAGGCAGT TATACCCCTT GTGATCGTGT CAGCCCTGAC TTTTATCTGT 2520  
 CTAGTGGTTC TTGTGGGTAT TCTCATCTAC TGGAGGAAAT GCTTCCAGAC TGCACACTTT 2580  
 TACTTAGAGG ACAGTACATC CCCTAGAGTT ATATCCACAC CTCCAACACC TATCTTTCCA 2640  
 85 ATTTTCAAGT ATGTCGGAGC AATTCCAATA AAGCACTTTC CAAAGCATGT TGCAGATTTA 2700  
 CATGCAAGTA GTGGGTTTAC TGAAGAATTT GAGACACTGA AAGAGTTTTC CCAGGAAGTG 2760  
 CAGAGCTGTA CTGTTGACTT AGGTATTACA GCAGACAGCT CCAACCACCC AGACAACAAG 2820



	CACAAGAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	2880
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	2940
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTC	3000
5	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	3060
	AAAGGAAGGA	GAAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	3120
	TTTCTGGTCA	CTCAGAAGAG	TGTGCAAGTG	CTTGCCCTATT	ATACTGTGAG	GAATTTTACT	3180
	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	3240
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	3300
10	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTCGTC	3360
	CACTGCAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	3420
	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAAACACAT	CCGTTCACAA	3480
	AGAAATTATT	TGGTACAAAC	TGAGGAGCAA	TATGTCTTCA	TTCATGATAC	ACTGGTTGAG	3540
	GCCATACTTA	GTAAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCATGCCTA	TGTTAATGCA	3600
	CTCCTCATT	CTGGACCAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GGGTCTCACT	3660
15	CTGTCACCCA	GGCTGGAGTG	CAGAGGCACA	ATCTCGGCTC	ACTGCAACCT	TCCTCTCCCT	3720
	GGCTTAACTG	ATCCTCCTAC	CTCAGCCTCC	CGAGTGGCTG	GGACTATACT	CCTGAGCCAG	3780
	TCAATATAC	AGCAGAGTGA	CTATTCTGCA	GCCCTAAAGC	AATGCAACAG	GGAAAAGAAT	3840
	CGAACTTCTT	CTATCATCCC	TGTGGAAGA	TCAAGGGTTG	GCATTTTCATC	CCTGAGTGGA	3900
	GAAGGCACAG	ACTACATCAA	TGCCTCCTAT	ATCATGGGCT	ATTACCAGAG	CAATGAATTC	3960
20	ATCATTACCC	AGCACCTCT	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	4020
	CATAATGCCC	AACCTGGTGGT	TATGATTCTT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	4080
	GTTTACTGGC	CAATAAAGA	TGAGCCTATA	AATTTGTGAGA	GCTTTAAGGT	CACTCTTATG	4140
	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTA	TAATTCAGGA	CTTTATCTTA	4200
25	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTCAGTGTCC	TAAATGGCCA	4260
	AATCCAGATA	GCCCCATTAG	TAAAACCTTT	GAACTTATAA	GTGTTATAAA	AGAAGAAGCT	4320
	GCCAAATAGG	ATGGGCCTAT	GATTGTTTCAT	GATGAGCATG	GAGGAGTGAC	GGCAGGAACT	4380
	TTCTGTGCTC	TGACAACCCT	TATGCACCAA	CTAGAAAAAG	AAAATTCCGT	GGATGTTTAC	4440
	CAGGTAGCCA	AGATGATCAA	TCTGATGAG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	4500
30	CAGTTTCTCT	ACAAAGTGAT	CCTCAGCCTT	GTGGGCACAA	GGCAGGAAGA	GAATCCATCC	4560
	ACCTCTCTGG	ACAGTAATGG	TGCAGCATTG	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	4620
	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	4680
	CTAAAATTAG	GCAGGAAAAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCC	ATCACCTGAC	4740
	AGTAACTTTC	ATGACATAGG	ATTCTGCCCG	CAAATTTATA	TCATTAACAA	TGTGTGCCCT	4800
35	TTTGCAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACTAAA	ATGATTGAAT	TTTACAGTAT	4860
	TTCTAAGAAT	GGAATTGTGG	TATTTTTTTC	TGTATTGATT	TTAACAGAAA	ATTTCAATTT	4920
	ATAGAGGTTA	GGAATTCCAA	ACTACAGAAA	ATGTTTGTTT	TTAGTGTCAA	ATTTTATAGT	4980
	GTATTGTAG	CAATTATCAG	GTTTGCTAGA	AATATAACTT	TTAATACAGT	AGCCTGTAAA	5040
	TAAACACTC	TTCCATATGA	TATTCAACAT	TTTACAACTG	CAGTATTCAC	CTAAAGTAGA	5100
40	AATAATCTGT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	5160
	ATAATTGTAG	ATTTTATAT	TTTACTACTG	AGTCAAGTTT	TCTAGTTCTG	TGTAATTGTT	5220
	TAGTTTAATG	ACGTAGTTCA	TTAGCTGGTC	TTACTCTACC	AGTTTTCTGA	CATTGTATTG	5280
	TGTTACCTAA	GTCATTAAC	TTGTTTCAGC	ATGTAATTTT	AACTTTTGTG	GAAAATAGAA	5340
	ATACCTTCAT	TTTGAAAGAA	GTTTTTATGA	GAATAACACC	TTACCAAACA	TTGTTCAAAT	5400
45	GGTTTTATC	CAAGGAATTG	CAAAAATAAA	TATAAATATT	GCCATTAAAA	AAAAAAAAAA	5460
	AAAAAAAAAA	AAAAAAAAAA	A				

Seq ID NO: 188 Protein sequence:  
Protein Accession #: EOS sequence

50	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNKKL	KFQGWDTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
55	FKASKITFW	GKCNMSSDGS	EHSLEGQKFP	LEMQUIYCFDA	DRFSSFEEAV	KGKGLRLALS	180
	ILFEVGTEN	LDFKAIIDGV	ESVSRFGKQA	ALDPFILNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVFCEVLTMQ	QSGYVMLMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGREIHEAV	CSSEPEVQA	DPENYTSLLV	TWERPRVVD	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNNLLP	NMSYVLQIVA	ICTNGLYGYK	SDQLIVDMPT	DNPELDLPE	420
60	LIGTEEIKE	EEEGKDIEEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTTHYNRI	GTKYNEAKTN	480
	RSPTRGSEFS	GKGDVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPHT	VEGTSASLND	540
	GSKTVLRSPH	MNLSGTAESL	NTVSITEYEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETITYD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NVWFPSSTDI	660
	TAQPDVGSGR	ESFLQNTYTE	IRVDESEKTT	KSFSAGPVMS	QGPSVTDLEM	PHYSTFAYFP	720
65	TEVTPHAFTP	SSRQQDLVST	VNVVYSQTTQ	PVYNEASNSS	HESRIGLAEG	LESEKKAVIP	780
	LVIVSALTFI	CLVVLVGILI	YWRKCFQTAH	FYLESTSPR	VISTPPTPIF	PISDDVGAIP	840
	IKHFPKHVAD	LHASSGFTEE	FETLKEFYQE	VQSCVDLGI	TADSSNHPDN	KHKNNRYINIV	900
	AYDHSRVKLA	QLAEKDGLT	DYINANYVDG	YNRPKAYIAA	QGPKLSTAED	FWRMIWEHNV	960
	EVIVMITNLV	EKGRRKCDQY	WPADGSEYEG	NFLVTQKSVQ	VLAYYTVRNF	TLRNTKIKKG	1020
70	SQKGRPSGRV	VTQYHYTQWP	DMGVPEYSLP	VLTFVRKAA	AKRHAVGPVV	VHCSAGVGR	1080
	GTIYVLDLML	QQIQHEGTVN	IFGFLKHRS	QRNYLVQTEE	QYVFIHDTLV	EAILSKETEV	1140
	LDSHIHAYVN	ALLIPGPAGK	TKLEKQFQGL	TLSPRLRCRG	TISAHCNLPL	PGLTDPPTSA	1200
	SRVAGTILLS	QSNIQSDYS	AALKQCNREK	NRTSSIIPE	RSRVGISSLS	GEGTDYINAS	1260
	YIMGYYSNE	FIITQHPLH	TIKDFWRMIW	DHNAQLVMI	PDQONMAEDE	FVYWPKNDEP	1320
75	INCESFKVTL	MAEEHKCLSN	EEKLIIQDFI	LEATQDDYVL	EVRHFQCPKW	PNPDSPISK	1380
	FELISVIKEE	AANRDGPMIV	HDEHGGVTAG	TFCALTLMH	QLEKENSVDV	YQVAKMINLM	1440
	RPGVFADIEQ	YQFLYKVILS	LVGTRQEENP	STSLDSNGAA	LPDGNIAESL	ESLV	

Seq ID NO: 189 DNA sequence  
Nucleic Acid Accession #: NM\_002820  
Coding sequence: 304..831

85	1	11	21	31	41	51	
	CCGGTTCGCA	AAGAAGCTGA	CTTCAGAGGG	GGAAACTTTC	TTCTTTTAGG	AGGCGGTTAG	60
	CCCTGTTCCA	CGAACCAGG	AGAACTGCTG	GCCAGATTAA	TTAGACATTG	CTATGGGAGA	120
	CGTGTAACA	CACTACTTAT	CATTGATGCA	TATATAAAC	CATTTTATTT	TCGCTATTAT	180



	TTCAGAGGAA	GCGCCTCTGA	TTTGTTCCTT	TTTTCCCTTT	TTGCTCTTTC	TGGCTGTGTG	240
	GTTTGGAGAA	AGCACAGTTG	GAGTAGCCGG	TTGCTAAATA	AGTCCCGAGC	GCGAGCGGAG	300
	ACGATGCAGC	GGAGACTGGT	TCAGCAGTGG	AGCGTCGCGG	TGTTCCCTGCT	GAGCTACGCG	360
5	GTGCCCTCCT	GCGGGCGCTC	GGTGGAGGGT	CTCAGCCGCC	GCCTCAAAAG	AGCTGTGTCT	420
	GAACATCAGC	TCCTCCATGA	CAAGGGGAAG	TCCATCCAAG	ATTTACGGCG	ACGATTCTTC	480
	CTTCACCATC	TGATCGCAGA	AATCCACACA	GCTGAAATCA	GAGCTACCTC	GGAGGTGTCC	540
	CCTAACTCCA	AGCCCTCTCC	CAACACAAAG	AACCACCCCG	TCCGATTGGG	GTCTGATGAT	600
	GAGGGCAGAT	ACCTAACTCA	GGAAGCTAAC	AAGGTGGAGA	CGTACAAAGA	GCAGCCGCTC	660
10	AAGACACCTG	GGAAGAAAAA	GAAAGGCAAG	CCCGGGAAAC	GCAAGGAGCA	GGAAAAGAAA	720
	AAACGGCGAA	CTCGCTCTGC	CTGGTTAGAC	TCTGGAGTGA	CTGGGAGTGG	GCTAGAAGGG	780
	GACCACCTGT	CTGACACCTC	CACAACGTCG	CTGGAGCTCG	ATTACCGGTA	ACAGGCTTCT	840
	CTGGCCCGTA	GCCTCAGCGG	GGTGCTCTCA	GCTGGGTTTT	GGAGCCTCCC	TTCTGCCTTG	900
	GCTTGACAA	ACCTAGAATT	TTCTCCCTTT	ATGTATCTCT	ATCGATTGTG	TAGCAATTGA	960
	CAGAGAATAA	CTCAGAATAT	TGTCTGCCTT	AAAGCAGTAC	CCCCCTACCA	CACACACCCC	1020
15	TGTCCTCCAG	CACCATAGAG	AGGCGCTAGA	GCCCATTCCCT	CTTTCTCCAC	CGTCACCCAA	1080
	CATCAATCCT	TTACCACTCT	ACCAATAAAT	TTCATATTCA	AGCTTCAGAA	GCTAGTGACC	1140
	ATCTTCATAA	TTTGCTGGAG	AAGTGTATTT	CTTCCCTTAA	CTCTCACACC	TGGGCAAACT	1200
	TTCTTCAGTG	TTTTTTCATTT	CTTACGTTCT	TTCATTCTAA	GGGAGAATAT	AGAAGCATTT	1260
20	GATATTATCT	ACAAACACTG	CAGAACAGCA	TCATGTCATA	AACGATTCTG	AGCCATTAC	1320
	ACTTTTTATT	TAATTAATG	TATTTAATTA	AATCTCAAAT	TTATTTTAAT	GTAAAGAACT	1380
	TAAATTATGT	TTTAAACACA	TGCCTTAAAT	TTGTTTAATT	AAATTTAACT	CTGGTTTCTA	1440
	CCAGCTCATA	CAAAATAAAT	GGTTTCTGAA	AATGTTTAAG	TATTAACCTA	CAAGGATATA	1500
	GGTTTTTCTC	ATGTATCTTT	TTGTTTCTTG	GCAAGATGAA	ATAATTTTTC	TAGGGTAATG	1560
25	CCGTAGGAAA	AATAAACTT	CACATTTAAA	AAAAA			

Seq ID NO: 190 Protein sequence:  
Protein Accession #: NP\_002811

30	1	11	21	31	41	51	
	MQRRLVQQWS	VAVFLLSYAV	PSCGRSVEGL	SRRLKRAVSE	HQLLHDKGKS	IQDLRRRFFL	60
	HHLIAEIHHTA	EIRATSEVSP	NSKPSPTNKN	HPVRFSGDDE	GRYLTQETNK	VETYEQPLK	120
35	TPGKKKKGKP	GKRKEQEKKK	RRTRSALWDS	GVTGSGLEGD	HLSDTSTTSL	ELDSR	

Seq ID NO: 191 DNA sequence  
Nucleic Acid Accession #: XM\_059328  
Coding sequence: 52..1023

40	1	11	21	31	41	51	
	GGGCTGTCCG	GCCCACTCCC	CTGGGAGCGC	GAGCGGTGGA	CCCAGGCGGC	CATGTCCCGC	60
	CCTCGCATGC	GCCTGGTGGT	CACCGCGGAC	GACTTTGGTT	ACTGCCCGCG	ACGCGATGAG	120
45	GGTATCGTGG	AGGCCTTTCT	GGCCGGGGCT	GTGACCAGCG	TGTCCCTGCT	GGTCAACGGT	180
	GCGGCCACGG	AGAGCGCGGC	GGAGCTGGCC	CGCAGGCACA	GCATCCCCAC	GGGCTCCAC	240
	GCCAACCTGT	CCGAGGGCCG	CCCCGTGGGT	CCGCCCCGCC	GTGGCGCCTC	ATCGCTGCTC	300
	GGCCCGGAAG	GCTTCTTCCT	TGGCAAGATG	GGATTCCGGG	AGGCGGTGGC	GGCCGGAGAC	360
50	GTGGATTTGC	CTCAGGTGCG	GGAGGAGCTC	GAGGCCCAAC	TAAGCTGCTT	CCGGGAGCTG	420
	CTGGGCAGGG	CCCCACGCA	CGCGGACGGG	CACCAGCACG	TGCACGTGCT	CCCAGGCGTG	480
	TGCCAGGTGT	TCGCCGAGGC	GCTGCAGGCC	TATGGGGTGC	GCTTTACGCG	ACTGCCGCTG	540
	GAGCGCGGTG	TGGGTGGCTG	CACTTGGCTG	GAGGCCCCCG	CGCGTGCCTT	CGCCTGCGCC	600
	GTGGAGCGCG	ACGCCCGGGC	CGCGTGGGCT	CCCTTCTCCC	GCCACGGCCT	GCGGTGGACA	660
55	GACGCCTTCG	TGGGCCTGAG	CACTTGCAGC	CGGCACATGT	CCGCTCACCG	CGTGTCCGGG	720
	GCCCTGGCGC	GGGTCCCTGA	AGGTACCCCTA	GCGGGCCACA	CCCTGACAGC	CGAGCTGATG	780
	GCGCACCCCG	GCTACCCCTG	TGTGCTCTCC	ACCGGCGGCT	GCGGTGAAGG	CCCCGACGCT	840
	TTCTCTTGCT	CTTGGGAGCG	GCTGCATGAG	CTGCGCGTCC	TCACCGCGCC	CACGCTGCGG	900
	GCCCAGCTTG	CCCAGGATGG	CGTGCAGCTT	TGCGCCCTCG	ACGACCTGGA	CTCCAAGAGG	960
60	CCAGGGGAGG	AGGTCCCTTG	TGAGAACCCCT	TCCTGGAACC	CTCCCTACTC		1020
	TGACCCCTTA	CAGACAACCA	AGCACTAATC	CCCTTAGTAC	CAAGAAAGGG	GAGCCAGGAT	1080
	TTAGTCCTGG	CCCAGCCCTG	AGCTGGGACC	TGGAGCACGA	TCTGTTGACT	TCCCTGGGTA	1140
	GGACACTGCC	ACCTCTGGGC	TCAGGTCTCT	ATGCCTCCAA	ATGGCATCTA	GAGTTTGAGC	1200
	AGCCTTCTTG	GCTGCAGGCA	GGCCTAGCCT	GTGGCAGCGG	GCTAGGGCCC	GCAGAGCATT	1260
65	TGGTGCCCTT	CCATGTTGCA	ATGCAAAACAC	CTTCACCACT	GGGGCAGTGG	GGAGAGATGG	1320
	CTATATTAAT	AAAATAACGT	GTGTCTTTC				

Seq ID NO: 192 Protein sequence:  
Protein Accession #: XP\_059328

70	1	11	21	31	41	51	
	MSRPRMRLVV	TADDFGYCPR	RDEGIVEAFL	AGAVTSVSLI	VNGAATESAA	ELARRHSIPT	60
	GLHANLSEGR	PVGPARRGAS	SLLGPEGFFL	GKMGFREAVA	AGDVLDPQVR	EELEAQLSCF	120
75	RELLGRAPTH	ADGHQHVHVL	PGVCQVFAEA	LQAYGVRFTR	LPLERGVGCG	TWLEAPARAF	180
	ACAVERDARA	AVGPFSRHGL	RWTDAPVGLS	TCGRHMSAHR	VSGALARVLE	GTLAGHTLTA	240
	ELMAHPGYPS	VPPTGGCGEG	PDAFSCSWER	LHELRLVLTAP	TLRAQLAQDG	VQLCALDDLD	300
	SKRPGEEVPC	EPTLEPFLEP	SLL				

Seq ID NO: 193 DNA sequence  
Nucleic Acid Accession #: NM\_005688.1  
Coding sequence: 126..4439

80	1	11	21	31	41	51	
	CCGGGCAGGT	GGCTCATGCT	CGGGAGCGTG	GTTGAGCGGC	TGGCGCGGTT	GTCCTGGAGC	60
85	AGGGGCGCAG	GAATTCTGAT	GTGAAACTAA	CAGTCTGTGA	GCCCTGGAAC	CTCCGCTCAG	120



	AGAAGATGAA	GGATATCGAC	ATAGGAAAAG	AGTATATCAT	CCCCAGTCCT	GGGTATAGAA	180
	GTGTGAGGGA	GAGAACCAGC	ACTTCTGGGA	CGCACAGAGA	CCGTGAAGAT	TCCAAGTTCA	240
	GGAGAACTCG	ACCGTTGGAA	TGCCAAGATG	CCTTGGAAAC	AGCAGCCCGA	GCCGAGGGCC	300
5	TCTCTCTTGA	TGCTTCCATG	CATTCTCAGC	TCAGAATCCT	GGATGAGGAG	CATCCCAAGG	360
	GAAAGTACCA	TCATGGCTTG	AGTGCTCTGA	AGCCCATCCG	GACTACTTCC	AAACACCAGC	420
	ACCCAGTGGA	CAATGCTGGG	CTTTTTTCT	GTATGACTTT	TTCGTGGCTT	TCTTCTCTGG	480
	CCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
10	AAGTTGGGCC	AGACGCTGCT	TCCCTGCCAA	GGGTGTGTGT	GATCTTCTGC	CGCACCAGGC	660
	TCATCCTGTC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAGCCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTTAGT	GCTGGGCTTC	CTCCTGACGG	AAATCGTGCG	GTCTTGGTCG	CTTGCACTGA	840
	CTTGGGCATT	GAATTACCGA	ACCGGTGTCC	GCTTGCGGGG	GGCCATCCTA	ACCATGGCAT	900
	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
15	TTTGCTCCAA	CGATGGGCAG	AGAATGTTTG	AGGCAGCAGC	CGTTGGCAGC	CTGCTGGCTG	1020
	GAGGACCCGT	TGTTGCCATC	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCCTGGG	ATCAGCTGTT	TTTATCCTCT	TTTACCCAGC	AATGATGTTT	GCATCACGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CCGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTTCTC	1260
20	AGAGTGTTCA	AAAAATCCGC	GAGGAGGAGC	GTCCGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGTATCAC	TGTGGGTGTG	GCTCCGATGG	TGGTGATGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTTCATAT	GACCCCTGGGC	TTCCGATCTGA	CAGCAGCACA	GGCTTTCACA	GTGGTGACAG	1440
	TCTTCAATTTC	CATGACTTTT	GCTTTGAAAG	TAACACCGTT	TTCAGTAAAG	TCCCTCTCAG	1500
25	AAGCCTCAGT	GGCTGTTGAC	AGATTTAAGA	GTTTGTCTCT	AATGGAAGAG	GTTTACATGA	1560
	TAAAGAACAA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAAAAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCACTCCAGT	ATCCAGAACT	CGCCCAAGCT	GACCCCCAAA	ATGAAAAAAG	1680
	ACAAGAGGGC	TTCCAGGGGC	AAGAAAGAGA	AGGTGAGGCA	GCTGCAGCGC	ACTGAGCATC	1740
	AGGCGGTGCT	GGCAGAGCAG	AAAGGCCACC	TCCTCCTGGA	CAGTGACGAG	CGGCCAGTGC	1800
30	CCGAAGAGGA	AGAAGGCAAG	CACATCCACC	TGGGCCACCT	GCGCTTACAG	AGGACACTGC	1860
	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AACTGGTTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCATT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACCTTC	GCTTATGTGG	CCCAGCAGGC	CTGGATCCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
35	ACAGCTGCTG	CCTGAGGCCT	GACCTGGCCA	TTCTTCCCAG	CAGCGACCTG	ACGGAGATTG	2160
	GAGAGCGAGG	AGCCAACCTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCCGGCCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTT	CAGTGCCCTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGATCTTC	ATGAAAGAGG	2400
40	GCTGTATTAC	GGAAAGAGGC	ACCATGAGG	AACTGATGAA	TTTAAATGGT	GACTATGCTA	2460
	CCATTTTTAA	TAACCTGTTG	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCAGTGG	TTACAGAAAG	AAGTCACAAG	ACAAGGGTCC	TAAAACAGGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCAGTGCC	CTGGTCAGTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
45	TCCTGGTTAT	TATGGCCCTT	TTCATGCTGA	ATGTAGGCAG	CACCGCCTTC	AGCACCTGGT	2760
	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAAACACCAC	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCCTGA	AAGCCATTCT	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGCG	AGCTTCCCTC	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTCAAGGCC	3000
50	CTATGAAGTT	TTTTGACACG	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
	TGGATGAAGT	TGACGTGCGG	CTGCCGTTCC	AGGCCGAGAT	GTTCATCCAG	AACGTTATCC	3120
	TGGTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCCT	GTGGCAGTGG	3180
	GGCCCCCTGT	GATCCTCTTT	TCAGTCTCTC	ACATGCTGAC	CAGGGTCCCT	ATTCGGGAGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCAGTCAC	CTTTCTCTCT	CCACATCAGC	TCCAGCATAC	3300
55	AGGGCCTTGC	CACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTCCTGCAC	AGATAACAGG	3360
	AGCTGCTGGA	TGACAACCAA	GCTCCTTTTT	TTTTGTTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACCAC	CACGGGGCTG	ATGATCGTTC	3480
	TTATGCACGG	GCAGATTCCC	CCAGCCTATG	CGGGTCTCGC	CATCTCTTAT	GCTGTCCAGT	3540
	TAACGGGGCT	GTTCCAGTTT	ACGGTCAGAG	TGGCATCTGA	GACAGAAGCT	CGATTACCTT	3600
60	CGGTGGAGAG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAAGCACCT	GCCAGAATTA	3660
	AGAACAAGGC	TCCCTCCCTT	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGCAG	3720
	AGATGAGGTA	CCGAGAAAAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAAC	3780
	CTAAAGAGAA	GATTGGCATT	GTGGGGCGGA	CAGGATCAGG	GAAGTCCTCG	CTGGGGATGG	3840
	CCCTCTTCCG	TCTGGTGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAAATCA	3900
65	GTGATATTGG	CCTTGCCGAC	CTCCGAAGCA	AACTCTCTAT	CATTCTCTAA	GAGCCGGTGC	3960
	TGTTACAGTG	CACTGTGAGA	TCAAATTTGG	ACCCCTTCAA	CCAGTACACT	GAAGACCAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TTGAATCTGA	AGTGATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
70	CCATGGACAC	AGAGACAGAC	TTATTGATTC	AAGAGACCAT	CCGAGAAGCA	TTTGCAAGCT	4260
	GTACCATGCT	GACCATTGCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCC	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTTCTGT	TCCAACGACA	4380
	GTTCCCGATT	CTATGCCATG	TTTGCTGCTG	CAGAGAACAA	GGTCGCTGTC	AAGGGCTGAC	4440
	TCCTCCCTGT	TGACGAAGTC	TCCTTTCTTT	AGAGCATTGC	CATTCCCTGC	CTGGGGCGGG	4500
75	CCCCCTCATG	CGTCCCTCTA	CCGAAACCTT	GCCTTTCTCG	ATTTTATCTT	TCGCACAGCA	4560
	GTTCCCGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTTGATTT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTCTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
	TCTATATATA	ATTCTGTACA	TAGCCTATTA	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
80	TATTAAAATA	AGCACTGTGC	TAATAACAGT	GCATATTCTT	TTCTATCATT	TTTGTACAGT	4860
	TTGCTGTACT	AGAGATCTGG	TTTTGTCTAT	AGACTGTAGG	AAGAGTAGCA	TTTCATTCTT	4920
	CTCTAGCTGG	TGGTTTTCAG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCTCTGCCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
	GAGACGGGTG	GGCGGCTGGA	GACCATGCAG	AGCGCCGTGA	GTTCTCAGGG	CTCCTGCCCT	5100
85	CTGTCTGTGT	GTCACTTACT	GTTTCTGTCA	GTGAGCAGC	GGGGCGAAGC	CCAGGCCCTT	5160
	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCTCT	TCTTCTTTTT	GCTGTTGTTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTTCT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340



GTTGGTTCCA AGCCCTGGAG CCAACTGCTG CTTTTTGAGG TGGCACTTTT TCATTTCCTT 5400  
 ATTCCCACAC CTCCACAGTT CAGTGGCAGG GCTCAGGATT TCGTGGGTCT GTTTTCCTTT 5460  
 CTCACCGCAG TCGTCGCACA GTCTCTCTCT CTCTCTCCCC TCAAAGTCTG CAACTTTAAG 5520  
 CAGCTCTTGC TAATCAGTGT CTCACACTGG CGTAGAAGTT TTTGTACTGT AAAGAGACCT 5580  
 ACCTCAGGTT GCTGGTTGCT GTGTGGTTTG GTGTGTTCCC GCAAACCCCT TTTGTGCTGT 5640  
 GGGGCTGGTA GCTCAGGTGG GCGTGGTCAC TGCTGTCATC AGTTGAATGG TCAGCGTTGC 5700  
 ATGTCGTGAC CAACTAGACA TTCTGTGCGC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760  
 CAAAAATCTG AAAATGTGAA TAAATATTAT TTGGATTTTG TAAAAAAA AAAAAAAA 5820  
 AAAAAAAA AAAAAAAA

Seq ID NO: 194 Protein sequence:  
 Protein Accession #: NP\_005679.1

1 11 21 31 41 51  
 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRR TRPLECQDAL ETAARAEGLS 60  
 LDASMHSQLR ILDEEHPKPK YHHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLAR 120  
 VAHKKGELSM EDVWSLSKHE SSDVNCRLLE RLWQELNEV GPDAASLRRV VWIFCRTRI 180  
 LSIVCLMITQ LAGFSGPAFM VKHLLLEYTQA TESNLQYSL LVLGLLLTEI VRSWSLALTW 240  
 ALNYRTGVRL RGAILTMAFK KILKLNKIKE KSLGELINIC SNDGQRMFEA AAVGSLLAGG 300  
 PVVAILGMIY NVIILGPTGF LGSVFIIFY PAMMFASRLT AYFRRKCVAA TDERVQKMNE 360  
 VLTYYIKFIK YAWVKAQS QVKIREEERR ILEKAGYFQG ITVGVAPIV VIASVVTFSV 420  
 HMTLGFDLTA QAFTVTVTF NSMTFALKVT PFSVKSLESEA SVAVDRFKSL FLMEEVHMIK 480  
 NKPASPHIKI EMKNATLAWD SSHSSIQNSP KLTPKMKKDK RASRGKKEKV RQLQRTEHQA 540  
 VLAEQKGHL LDDSDERPSPE EEEGKHIHLG HLRLQRTLHS IDLEIQEGKL VGICGSVGS 600  
 KTSLSAILG QMTLLEGSIA ISGTFAYVAQ QAWILNATLR DNILFGKEYD EERYNSVLNS 660  
 CCLRPDLAIL PSSDLTEIGE RGANLSGGQR QRISLARALY SDRSIYILDD PLSALDAHVG 720  
 NHIFNSAIRK HLKSKTVLFV THQLQYLVDC DEVIFMKEGC ITERGTHEEL MNLNGDYATI 780  
 FNNLLLGETP PVEINSKKEK SGSQKKSQDK GPKTGSVKKE KAVKPEEQQL VQLEEKQGS 840  
 VPWSVYGVYI QAAGGPLAFL VIMALFMLNV GSTAFSTWWL SYWIKQSGSN TTVTRGNETS 900  
 VSDSMKDNPH MQYYASIIYAL SMAVMLILKA IRGVVFVKGT LRASSRLHDE LFRRLRSPM 960  
 KFFDTPTTGR ILNRFSKMDM EVDVRLPFQA EMFIQNVILV FFCVGMIAV FFWFLVAVGP 1020  
 LVILFSVLHI VSRVLIRELK RLDNITSPF LSHITSSIQG LATIHAYNKG QEFLHRYQEL 1080  
 LDDNQAPFFL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HGQIPPAYAG LAISYAVQLT 1140  
 GLFQFTVRLA SETEARFTSV ERINHYIKTL SLEAPARIKN KAPSPDWPQE GEVTFENAEM 1200  
 RYRENLPVLV KKVSTFIKPK EKIGIVGR TG SGKSSLGMA FRLVELSGGC IKIDGVRISD 1260  
 IGLADLRSL SIIPQEPVLF SGTVRSNLD FNQYTEDQIW DALERTHMEK CIAQLPLKLE 1320  
 SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLLEQE TIREAFADCT 1380  
 MLTIAHRLHT VLGS DRIMVL AQGQVVEFDT PSVLLSNDSS RFYAMFAAAE NKVAVKG

Seq ID NO: 195 DNA sequence  
 Nucleic Acid Accession #: NM\_006470  
 Coding sequence: 228..1922

1 11 21 31 41 51  
 GCTGTCCTGA GCCTGAGTAC TCTAGCTGCC TTGTCGCCAT CGCATCTGGC TGCCATCCAG 60  
 CGCCAGCACA CAGTAATGAG TGGCCGAGCT TCCTCTGGGA GGGAGGAAAC AGTTAAATC 120  
 TTGCAGCAGC TGCAATCATC TAGGCGTGGT TCTCTGTCT GACTTGGGCT GCACAGATCC 180  
 TGGGCCAAGG GACAGAAGAA AGACAGCCTA GGAGCAGAGC CTCCCAGATG GCTGAGTTGG 240  
 ATCTAATGGC TCCAGGGCCA CTGCCCAGGG CCACTGCTCA GCCCCAGCC CCTCTCAGCC 300  
 CAGACTCTGG GTCAACCCAGC CCAGATTCTG GGTCAGCCAG CCCAGTGGAA GAAGAGGACG 360  
 TGGGCTCCTC GGAGAAGCTT GGCAGGGAGA CGGAGGAACA GGACAGCGAC TCTGCAGAGC 420  
 AGGGGGATCC TGCTGGTGAG GGGAAAGAGG TCCTGTGTGA CTTCTGCCTT GATGACACCA 480  
 GAAGAGTGAA GGCAGTGAAG TCCTGTCTAA CCTGCATGGT GAATTACTGT GAAGAGCACT 540  
 TGCAGCCGCA TCAGGTGAAC ATCAAATGTC AAAGCCACCT GCTGACCGAG CCAGTGAAGG 600  
 ACCACAATG GCGATACTGC CCTGCCCACC ACAGCCACT GTCTGCTTTC TGCTGCCCTG 660  
 ATCAGCAGTG CATCTGCCAG GACTGTGACC AGGAGCACAG TGGCCACACC ATAGTCTCCC 720  
 TGGATGCAGC CCGCAGGGAC AAGGAGGCTG AACTCCAGTG CACCCAGTTA GACTTGGAGC 780  
 GGAAACTCAA GTTGAATGAA AATGCCATCT CCAGGCTCCA GGCTAACCAG AAGTCTGTTC 840  
 TGGTGTGCGT GTCAGAGGTC AAAGCGGTGG CTGAAATGCA GTTTGGGGAA CTCCTTGCTG 900  
 CTGTGAGGAA GGCCAGGCC AATGTGATGC TCTTCTTAGA GGAGAAGGAG CAAGCTGCGC 960  
 TGAGCCAGGC CAACGGTATC AAGGCCACC TGGAGTACAG GAGTGCCGAG ATGGAGAAGA 1020  
 GCAAGCAGGA GCTGGAGAGG ATGGCCGCCA TCAGCAACAC TGTCCAGTTC TTGGAGGAGT 1080  
 ACTGCAAGTT TAAGAACACT GAAGACATCA CCTTCCCTAG TGTTCACGTA GGGCTGAAGG 1140  
 ATAACTCTC GGGCATCCGC AAAGTTATCA CGGAATCCAC TGTACACTTA ATCCAGTTGC 1200  
 TGGAGAATA TAAGAAAAG CTCCAGGAGT TTTCCAAGGA AGAGGAGTAT GACATCAGAA 1260  
 CTCAAGTGTC TGCCGTGTGT CAGCGCAAAT ATTGGACTTC CAAACCTGAG CCCAGCACCA 1320  
 GGGAAAGATT CCTCCAATAT GCGTATGACA TCACGTTTGA CCCGACACA GCACACAAGT 1380  
 ATCTCCGGCT GCAGGAGGAG AACCAGCAAG TCACCAACAC CACGCCCTGG GAGCATCCCT 1440  
 ACCCGACCT CCCAGCAGG TTCCTGCACT GCGGCGAGGT GCTGTCCAG CAGAGTCTGT 1500  
 ACCTGCACAG GTACTATTTT GAGGTGGAGA TCTTCGGGGC AGGCACCTAT GTTGGCCTGA 1560  
 CCTGCAAAGG CATCGACCGG AAAGGGGAGG AGCGCAACAG TTGCATTTCC GGAAACAACT 1620  
 TCTCCTGGAG CCTCCAATGG AACGGGAAGG AGTTCACGGC CTGGTACAGT GACATGGAGA 1680  
 CCCCCTCAA AGCTGGCCCT TTCCGGAGGC TCGGGGTCTA TATCGACTTC CCGGAGGGA 1740  
 TCCTTTCTCT CTATGGCGTA GAGTATGATA CCATGACTCT GGTTACAAAG TTTGCCTGCA 1800  
 AATTTTCAGA ACCAGTCTAT GCTGCCTTCT GGCTTTCCAA GAAGGAAAAC GCCATCCGGA 1860  
 TTGTAGATCT GGGAGAGGAA CCCGAGAAGC CAGCACCGTC CTTGGGGGTG ACTGCTCCCT 1920  
 AGACTCCAGG AGCCATATCC CAGACCTTTG CCAGCTACAG TGATGGGATT TGCAATTTAG 1980  
 GGTGATTTGT GGGCAGAAAT AACTGCTGAT GGTAGCTGGC TTTTGAAATC CTATGGGGTC 2040  
 TCTGAATGAA AACATTCTCC AGCTGCTCTC TTTTGCTCCA TATGGTGCTG TTCTCTATGT 2100  
 GTTTGCAGTA ATTCTTTTTT TTTTTTTTGA GACGAGTCT CGCACTGTTG CCCAGGCTGG 2160  
 AGAGCAGTGG CGCATCTTG GCTCACTGCA AGCTCCGCT CCCGAGTTCA AGCAATTCTC 2220  
 CTGCCTCAGC CTCCCGAGTA GCTGGGATTA CAGGTGCCTG CCACCACACC CAGCTAATGT 2280  
 TTTGTATTTT TAGTAGAGAT GGGGTTTCAC CATGTTGGCC AGGCAGATCT CAAACTCCTG 2340



ACCTCGTGAT GCACCCACCT CGGCCTCCCA AAGTGCTGGG ATTACATGCG TGAGCCACTG 2400  
 CGCCCTGCCT GTTTGTAGTA ATTTTATAGG ACCAAATCTC CCTCATCTTC TAGTGCCATT 2460  
 CTCCTCTCTG TTCAGGTAAA TGTCACACTG TGCCCAGAAAT GGATGACCAG GAACCTTAAA 2520  
 GAGTGGCTGA AAAGATTGCA GAGTTATCAT AATAAATTGC TAACTTGCGT

Seq ID NO: 196 Protein sequence:  
 Protein Accession #: NP\_006461

1 11 21 31 41 51  
 MAELDLMAPG PLPRATAQPP APLSPDSGSP SPDSGSASPV EEEDVGSSEK LGRETEEQDS 60  
 DSAEQGDPAG EGKEVLCDFC LDDTRRVKAV KSCLTCMVNY CEEHLQPHQV NIKLQSHLLT 120  
 EPVKDHNWRY CPAHHSPLSA FCCPDQQCIC QDCCQEHSQH TIVSLDAARR DKEAELQCTQ 180  
 15 LDLERKLLKN ENAISRLQAN QKSVLVSVSE VKAVAEMQFG ELLAAVRKAQ ANVMLFLEEK 240  
 EQAALSQANG IKAHLEYRSA EMEKSKQELE RMAAISNTVQ FLEEYCKFKN TEDITFPSVY 300  
 VGLKDKLSGI RKVITESTVH LIQLLENYKK KLQEFSSKEE YDIRTQVSAV VQRKYWTSKP 360  
 EPSTREQFLQ YAYDITFDPD TAHKYLRLQE ENRKVTNTTP WEHPYPDLPS RFLHWRQVLS 420  
 20 QQSLYLHRY YFEVEIFGAGT YVGLTCKGID RKGEERNSCI SGNNFSWSLQ WNGKEFTAWY 480  
 SDMETPLKAG PFRRLGVYID FPGGILSFYV VEYDTMTLVH KFAKCFSEPV YAAFWLSKKE 540  
 NAIRIVDLGE EPEKPAPSLG VTAP

Seq ID NO: 197 DNA sequence  
 Nucleic Acid Accession #: NM\_004316  
 Coding sequence: 433-1149

1 11 21 31 41 51  
 CCCGAGACCC GCGCAAGAG AGCGCAGCCT TAGTAGGAGA GGAACGCGAG ACGCGGCAGA 60  
 30 GCGCGTTCAG CACTGACTTT TGCTGCTGCT TCTGCTTTTT TTTTCTTAG AAACAAGAAG 120  
 GCGCCAGCGG CAGCCTCACA CGCGAGCGCC ACGCGAGGCT CCCGAGCCA ACCCGCGAAG 180  
 GGAGGAGGGG AGGGAGGAGG AGGCGGCGTG CAGGGAGGAG AAAAAGCATT TTCACCTTTT 240  
 TTGCTCCAC TCTAAGAAGT CTCCCGGGG TTTTGTATAT ATTTTAAAC TTCGTCAGG 300  
 35 GCTCCCGCTT CATATTTCC TTTCTTTCC TCTCTGTTCC TGCACCAAG TTCTCTCTGT 360  
 GTCCCCCTCG CGGGCCCCGC ACCTCGCGTC CCGGATCGCT CTGATTCCGC GACTCCTTGG 420  
 CCGCCGCTGC GCATGGAAG CTCTGCAAG ATGGAGAGCG GCGGCGCCG CCAGCAGCCC 480  
 CAGCCGCGAG CCCAGCAGCC CTTCCTGCCG CCCGAGCCT GTTCTTTGC CACGGCCGCA 540  
 GCGCGGCGG CCGCAGCCGC CGCAGCGGCG GCGCAGAGCG CGCAGCAGCA GCAGCAGCAG 600  
 CAGCAGCAG AGCAGCAGCA GCAGGCGCG CAGCTGAGAC CGGCGGCCG CGGCCAGCCC 660  
 40 TCAGGGGGCG GTCAAGAGT AGCGCCCAAG CAAGTCAAGC GACAGCGCTC GTCTTCGCCC 720  
 GAACTGATGC GCTGCAACG CCGGCTCAAC TTCAGCGCT TGGGCTACAG CCTGCCGAG 780  
 CAGCAGCCG CCGCCGTGGC GCGCCGCAAC GAGCGCGAGC GCAACCGCGT CAAGTTGGTC 840  
 AACCTGGGCT TTGCCACCT TCGGGAGCAC GTCCCAACG GCGCGGCCA CAAGAAGATG 900  
 45 AGTAAGGTGG AGACACTGCG CTCGGCGGTC GAGTACATCC GCGCGCTGCA GCAGCTGCTG 960  
 GACGAGCATG ACGCGGTGAG CGCCGCTTC CAGGAGGCG TCCTGTCGCC CACCATCTCC 1020  
 CCCAACTACT CCAACGACTT GAACTCCATG CCGGCTCGC CGGTCTCATC CTACTCGTCG 1080  
 GACGAGGGCT CTTACGACCC GCTCAGCCCC GAGGAGCAGG AGCTTCTCGA CTTACCAAC 1140  
 TGGTTCTGAG GGGCTCGGCC TGGTCAGGCC CTGGTGCGAA TGGACTTTGG AAGCAGGGTG 1200  
 50 ATCGCACAA CTGCATCTT AGTGCTTTCT TGTCAGTGCG GTTGGGAGGG GGAGAAAAGG 1260  
 AAAAGAAAAA AAAAGAAAG GAAGAAGAA AGAGAAGAAG AAAAAACGA AAACAGTCAA 1320  
 CCAACCCCAT CGCCAACATA GCGAGGCATG CCTGAGAGAC ATGGCTTTCA GAAAACGGGA 1380  
 AGCGCTCAGA ACAGTATCTT TGCACTCAA TCATTACAG AGATATGAAG AGCAACTGGG 1440  
 ACCTGAGTCA ATGCGCAAAA TGCACTTGT TGCAAAAAG AGTGGGCTCC TGGCAGAAGG 1500  
 55 GAGCAGCACA CGCGTTATAG TAACTCCCAT CACCTCTAAC ACGCACAGCT GAAAGTTCTT 1560  
 GCTCGGGTCC CTTACCTCC CCGCCCTTTC TTAGAGTGCA GTTCTTAGCC CTCTAGAAAC 1620  
 GAGTTGGTGT CTTTC

Seq ID NO: 198 Protein sequence:  
 Protein Accession #: NP\_004307

1 11 21 31 41 51  
 MESSAKMESG GAGQQPQPQP QPFLPPAAC FFATAAAAAA AAAAAAQSA QQQQQQQQQQ 60  
 65 QQQQAPQLRP AADGQPSGGG HKSAPQVKR QRSSPELMR CKRRLNFSGF GYSLPQQQPA 120  
 AVARRNERER NRVKLVLNLF ATLREHVPNG AANKKMSKVE TLRSAYEYIR ALQQLLDEHD 180  
 AVSAAFQAGV LSPTISPNYS NDLNSMAGSP VSSYSSDEGS YDPLSPEEQE LLDFTNWF

Seq ID NO: 199 DNA sequence  
 Nucleic Acid Accession #: NM\_007015  
 Coding sequence: 1-1005

1 11 21 31 41 51  
 ATGACAGAGA ACTCCGACAA AGTTCCCATT GCCCTGGTGG GACCTGATGA CGTGGAATTC 60  
 75 TGCAGCCCCC CGGCGTACGC TACGCTGACG GTGAAGCCCT CCAGCCCCGC GCGGCTGCTC 120  
 AAGGTGGGAG CCGTGGTCTT CATTTGCGGA CACTGTCTGC TGCTCTTGG GCCATCGGG 180  
 GCCTTCTACT TCTGGAAGGG GAGCGACAT GACATTTACA ATGTCCATTA CACCATGAGT 240  
 ATCAATGGGA AACTACAAGA TGGGTCAATG GAAATAGACG CTGGGAACAA CTTGGAGACC 300  
 80 TTTAAATGG GAAGTGGAGC TGAAGAAGCA ATTGCAGTTA ATGATTCCA GAATGGCATC 360  
 ACAGGAATTC GTTTTGCTGG AGGAGAGAAG TGCTACATTA AAGCGCAAGT GAAGGCTCGT 420  
 ATTCTGAGG TGGGCGCCGT GACCAAACAT AGCATCTCCT CCAACTGGA AGGCAAGATC 480  
 ATGCCAGTCA AATATGAAGA AAATCTCTT ATCTGGGTGG CTGTAGATCA GCCTGTGAAG 540  
 85 GACAACAGCT TCTTGAGTTC TAAGGTGTTA GAACCTGCG GTGACCTTC TATTTTCTGG 600  
 CTTAAACCAA CCTATCCAAA AGAAATCCAG AGGGAAAGAA GAGAAGTGGT AAGAAAATT 660  
 GTTCCAAC TAACAAAAAG ACCACACAGT GGACCACGGA GCAACCCAGG CGCTGGAAGA 720  
 CTGAATAATG AAACCAGACC CAGTGTTCAG GAGGACTCAC AAGCCTTCAA TCCTGATAAT 780



5  
10  
CCTTATCATC AGCAGGAAGG GGAAAGCATG ACATTGACCC CTAGACTGGA TCACGAAGGA 840  
ATCTGTTGTA TAGAATGTAG GCGGAGCTAC ACCCACTGCC AGAAGATCTG TGAACCCCTG 900  
GGGGGCTATT ACCCATGGCC TTATAATTAT CAAGGCTGCC GTTCGGCCTG CAGAGTCATC 960  
ATGCCATGTA GCTGGTGGGT GGCCCGTATC TTGGGCATGG TGTGAAATCA CTTCATATAT 1020  
CACGTGCTGT AAAATAAGAA CTAGCTGAAG AGACAACCAA AGAAGCATTA AGGCAGGTTG 1080  
ATGCTGATGG GACCATAAAA TATTTTACCA CGCAGCCTGA GCGGTTATTC TTGACACTCT 1140  
TAACAGAATT TTTTAAATCG TTTTCCAGAA CTTTAGTATA TGCAAATGCA CTGAAAGGGT 1200  
AGTTCAAGTC TAAATGCGA TAACCCCGTT ATTTGTTATT TTTTATTTGC ATTGATTGTC 1260  
CATAAGTCTT CCCTTGCTTG CATCTTCCAA AGCTATTTTC AAATAAACAC GAAAATTTAC 1320  
AGTTTGCC

Seq ID NO: 200 Protein sequence:  
Protein Accession #: NP\_008946

15  
20  
25  
1 11 21 31 41 51  
| | | | | |  
MTENSDKVPI ALVGPDDVEF CSPPAYATLT VKPSSPARLL KVGAVVLISG AVLLLFGAIG 60  
AFYFWKGS DS HIYNVHYTMS INGKLQD GSM EIDAGNNLET FKMGS GAEEA IAVNDFQNGI 120  
TGIRFAGGEK CYIKAQVKAR IPEVGAVTKQ SISSKLEKGI MPVKYEENSL IWVAVDQPVK 180  
DNSFLSSKVL ELCDLPIFW LKPTYPKIEIQ RERREVVRKI VPTTTKRPHS GPRSNPGAGR 240  
LNNETRPSVQ EDSQAFNPDN PYHQEGESM TFDPRLDHEG ICCIECRRSY THCQKICEPL 300  
GGYYPWPYNY QGCRSACRVI MPCSWWVARI LGMV

Seq ID NO: 201 DNA sequence  
Nucleic Acid Accession #: NM\_000728.2  
Coding sequence: 112..495

30  
35  
40  
45  
50  
55  
60  
65  
1 11 21 31 41 51  
| | | | | |  
GTAATAAGAG CGGGGTCTCC GCGGGGAAGG CGCCACACAGC AGGTGTGGTG TTCATCCCGG 60  
GTGACCGGCG CGTTCGCGCT GCCCTGAAAC TCTAGTCGCC AGAGAGGCGG CATGGGTTTC 120  
CGGAAGTTCT CCCCCTTCCT GGCTCTCAGT ATCTTGGTCC TGTACCAGGC GGGCAGCCTC 180  
CAGGCGGCGC CATTCAAGGTG TGCCCTGGAG AGCAGCCCGAG ACCCGGCCAC ACTCAGTAAA 240  
GAGGACGCGC GCCTCCTGCT GGCTGCACTG GTGCAGGACT ATGTGCAGAT GAAGGCCAGT 300  
GAGCTGAAGC AGGAGCAGGA GACACAGGGC TCCAGCTCCG CTGCCCAGAA GAGAGCCTGC 360  
AACACTGCCA CCTGTGTGAC TCATCGGCTG GCAGGCTTGC TGAGCAGATC AGGGGGCATG 420  
GTGAAGAGCA ACTTCGTGCC CACCAATGTG GGTTCCAAAG CCTTTGGCAG GCGCCGCGAG 480  
GACCTTCAAG CCTGAGCAGA TGAATGACTC CAGGAAGAAG GTGTGTCCTA AATCCAATGA 540  
CATATCCTTA TAAGAGATTC ACTCAGAAGA CACATGTGGA GAAGGTGACA TGACAGAGGC 600  
AAGGAGGCAC AAGCCAAGGA AGTCTGTGTC TACCAGAAGC CAGAATCACA GAACAGTCTC 660  
TGGAAGAAGA GCAGCCCTGC TGACACCTAG AGTTTGGACT TCCAGCTTCC AGAACTGTGA 720  
GAGAATAATT TCTGTTGTTT TAAGCCACAA AGTTTGTGGT AATTTGTTAT GACAGCCCTA 780  
GGAAACTAAT ACAATACATT TTCAATTTATT TTGGGTAAAT GCCTTGGAGT GGGATTGCTG 840  
GGTTATTTGG AAAGTGTGTA TTTAACTCTG TAAGAAACTG CCAAATATT TTCTGAAGTG 900  
ACTGTACCAC TTCGCTTCTT TGCCAGCCAC ATATGAGAGC TCTAGTATTT CCACAATAG 960  
GTATGTAGCA GTATCTCATT GCTGTTTAA TTTGTATTTT CCAATGACT AATGACGTTG 1020  
AGCATCTATT TTACCATATG TTTATCACCT TTATTGAAGG GTCTGTTTAA ATCTTCTGCT 1080  
AAATTTTGTG TGGCTTGCTT GCTTTATTAG TGTGTAGTTT TTAGAGCTCT TTATATGTTG 1140  
TGGATGCAAG ATTGTTTTCA GATATATAGT TTGGAAACTT CCTTCCCTG AATCTGCGGA 1200  
TTGCTTTTTT ATTTTCTTAG CAGTGTCTCT CACAGAGAAA AAGTTGTAAT TTGAATAAGA 1260  
TCCAATTCAT CTTTTTTTTT CTTTTATGTA TTGTGCTTTT AGTTTCATGT TAAGAACTCT 1320  
TTGCCTAACT AAGGTCCCAA GGTCACAATA ACCTTATTCT ATACTTCTT GTAAAGTTT 1380  
TATAGTTTTA TATTTTATAT GTAGATTAGT GATCTATTTT GAGTTAATTT TTGTATAAGG 1440  
TGAGAGGTGT AGGTGAAAT TCATACCTGT GAATATAGAT ACCCAATTGT TTCAGTGCCA 1500  
TTTGTAAAAA AGACTGTTAT TTCACCATTT AATTGCCCCC GCACCTTTGT CAAAAAGCAA 1560  
CTGATCATAT TTGTGTGGGT ATATTTCTGG GTTCTCAATT CTGTCTCATT GATTGATTTG 1620  
ACCATTCTTT TGCCAATGTC ATACTGCCCT GATTAGTGTA GTGTTAAAGT GAATCTCAAA 1680  
ACCAGATAAT GTGGGTCTAC CAACATGTGT CATTCTTGTT CAAAAAGATT TTAGCTACAT 1740  
CTAAAATATT TTCTACATCT TTTATACAT TTAGAATCAG TGTGTTACTA TCTACAAAAT 1800  
TTCTGATGAG ATTTTAAATG GGATTGTGTT AAATCAGTGG GTTAATTTTG GGAGAATTAG 1860  
CATATTAATA ATATTAAGTC GTTCAATTCA TGAACACAAT ACATGTTTTT ACTTATTTAG 1920  
GTTTTCTCTG TTTTTTTTTT TTTAACAGTG TTCTCAGTTT TCAACAGAAA TATTCTACAC 1980  
ATATCTTGTT AGATTTTTAA CTATTTTATT TTTTGGTGCT AATGTAAATG GTACTTAAAC 2040  
ATTTTGTGTT TTAATTGTTT ATTGCTAGTA GATAGAAATA CAATATTTAA AATATTAGGA 2100  
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 202 Protein sequence:  
Protein Accession #: NP\_000719.1

70  
75  
1 11 21 31 41 51  
| | | | | |  
MGFRKFSFPL ALSILVLYQA GSLQAAPFRS ALESSPDPAT LSKEDARLLL AALVQDYVQM 60  
KASELKQEQE TQSSSAAQK RACNTATCVT HRLAGLLSRS GGMVKS NFVP TNVGSKAFGR 120  
RRRDLQA

Seq ID NO: 203 DNA sequence  
Nucleic Acid Accession #: NM\_001741  
Coding sequence: 71..496

80  
85  
1 11 21 31 41 51  
| | | | | |  
CTCTGGCTGG ACGCCGCCGC CGCCGCTGCC ACCGCTCTG ATCCAAGCCA CCTCCCGCCA 60  
GAGAGGTGTC ATGGGCTTCC AAAAGTTCTC CCCCTTCTG GCTCTCAGCA TCTTGGTCCT 120  
GTTGCAGGCA GGCAGCTCC ATGCAGCACC ATTCAAGTCT GCCCTGGAGA GCAGCCCGAGC 180  
AGACCCGGCC ACGCTCAGTG AGGACGAAGC GCGCCTCTG CTGGCTGCAC TGGTGCAGGA 240  
CTATGTGCAG ATGAAGGCCA GTGAGCTGGA GCAGGAGCAA GAGAGAGAGG GCTCCAGCCT 300



5 GGACAGCCCC AGATCTAAGC GGTGCGGTAA TCTGAGTACT TGCATGCTGG GCACATACAC 360  
 GCAGGACTTC AACAAAGTTT ACACGTTCCC CCAAAGTACA ATTGGGGTTG GAGCACCTGG 420  
 AAAGAAAAGG GATATGTCCA GCGACTTGGA GAGAGACCAT CGCCCTCATG TTAGCATGCC 480  
 CCAGAATGCC AACTAACTC CTCCCTTTCC TTCCTAATTT CCCTTCTTGC ATCCTTCCTA 540  
 TAACTTGATG CATGTGGTTT GGTTCCTCTC TGGTGGCTCT TTGGGCTGGT ATTGGTGGCT 600  
 TTCCTTGTGG CAGAGGATGT CTCAAACCTC AGATGGGAGG AAAGAGAGCA GGAATCACAG 660  
 GTTGAAGAG AATCACCTGG GAAAATACCA GAAAATGAGG GCCGCTTTGA GTCCCCCAGA 720  
 GATGTCATCA GAGCTCCTCT GTCCTGCTTC TGAATGTGCT GATCATTTGA GGAATAAAAT 780  
 10 TATTTTTCCC C

Seq ID NO: 204 Protein sequence:  
 Protein Accession #: NP\_001732

15 1 11 21 31 41 51  
 | | | | | |  
 MGFQKFSPFL ALSILVLLQA GSLHAAPFRS ALESSPADPA TLSEDEARLL LAALVQDYVQ 60  
 MKASELEQEQ EREGSSLDSP RSKRCGNLST CMLGTYTQDF NKFHTFPQTA IGVGAPGKKR 120  
 20 DMSSDLERDH RPHVSMPQNA N

Seq ID NO: 205 DNA sequence  
 Nucleic Acid Accession #: NM\_005361  
 Coding sequence: 1-945

25 1 11 21 31 41 51  
 | | | | | |  
 ATGCCTCTTG AGCAGAGGAG TCAGCACTGC AAGCCTGAAG AAGGCCTTGA GGCCCGAGGA 60  
 GAGGCCCTGG GCCTGGTGGG TGCGCAGGCT CCTGCTACTG AGGAGCAGCA GACCGCTTCT 120  
 30 TCCTCTTCTA CTCTAGTGGA AGTTACCTCG GGGGAGGTGC CTGCTGCCGA CTCACCGAGT 180  
 CCTCCCCACA GTCCTCAGGG AGCCTCCAGC TTCTCGACTA CCATCAACTA CACTCTTTGG 240  
 AGACAATCCG ATGAGGGCTC CAGCAACCAA GAAGAGGAGG GGCCAAGAAT GTTTCCCGAC 300  
 CTGGAGTCCG AGTTCCAAGC AGCAATCAGT AGGAAGATGG TTGAGTTGGT TCATTTTCTG 360  
 CTCCTCAAGT ATCGAGCCAG GGAGCCGGTC ACAAAGGCAG AAATGCTGGA GAGTGTCTCT 420  
 35 AGAAATTGCC AGGACTTCTT TCCCGTGATC TTCAGCAAAG CCTCCGAGTA CTTGCAGCTG 480  
 GTCTTTGGCA TCGAGGTGGT GGAAGTGGTC CCCATCAGCC ACTTGACAT CTTGTGACC 540  
 TGCCTGGGCC TCTCCTACGA TGGCCTGCTG GCGACAATC AGGTCATGCC CAAGACAGGC 600  
 CTCCTGATAA TCGTCTGGC CATAATCGCA ATAGAGGGCG ACTGTGCCCC TGAGGAGAAA 660  
 ATCTGGGAGG AGCTGAGTAT GTTGGAGGTG TTTGAGGGGA GGGAGGACAG TGTCTTCGCA 720  
 40 CATCCAGGA AGCTGCTCAT GCAAGATCTG GTGCAGGAAA ACTACCTGGA GTACCGGCAG 780  
 GTGCCCGGCA GTGATCCTGC ATGCTACGAG TTCCTGTGGG GTCCAAGGGC CCTCATTGAA 840  
 ACCAGCTATG TGAAAGTCCT GCACCATACT CTAAAGATCG GTGGAGAACC TCACATTTCC 900  
 TACCCACCCC TGCATGAACG GGCTTTGAGA GAGGGAGAAG AGTGA

Seq ID NO: 206 Protein sequence:  
 Protein Accession #: NP\_005352

45 1 11 21 31 41 51  
 | | | | | |  
 MPLEQRSQHC KPEEGLEARG EALGLVGAQA PATEEQQTAS SSSTLVEVTL GEVPAADSPS 60  
 50 PPHSPQGASS FSTTINYTLW RQSDDESSNQ EEEGPRMFPD LESEFQAAIS RKMVELVHFL 120  
 LLKYRAREPV TKAEMLESVL RNCQDFFPVI FSKASEYLQL VFGIEVVEVV PISHLYILVT 180  
 CLGLSYDGLL GDNQVMPKGT LLIIIVLAIIA IEGDCAPEEK IWEELSMLEV FEGREDSVFA 240  
 HPRKLLMQDL VQENYLEYRQ VPGSDPACYE FLWGPRALIE TSYVKVLHHT LKIGGEPHIS 300  
 55 YPPLHERALR EGEE

Seq ID NO: 207 DNA sequence  
 Nucleic Acid Accession #: NM\_021115  
 Coding sequence: 743-2893

60 1 11 21 31 41 51  
 | | | | | |  
 AAAGGAAGGG AGGGAGGGAG AAAGGAGAAG TTGGTTTAGA GGCCAGCCGG ACGAGCTTTG 60  
 GGCACCGCCC TTAGGAGGGC CACCCTCAGA GTCTGACAGC AGGTGAAGGT CCTAAATCTC 120  
 65 CCCAACTAA CTGGTGTCTT TTCTCCTCTT CCAAGATGCT CTTCCCGAGG GAGATGCTAG 180  
 CCTTTGGGT CCTTACCTCC TGCCCTCAGT AGCCCCGGAG AGAGGCAGTC CTGGCAAAGA 240  
 GCACCCCTGAA GAGAGAGTGG TAACAGCGCC CCCAGTTCC TCACAGTCGG CGGAAGTGCT 300  
 GGGCGAGCTG GTGCTGGATG GGACCGCACC CTCTGCACAT CACGACATCC CAGCCCTGTC 360  
 ACCGCTGCTT CCAGAGGAGG CCCGCCCAAA GCACGCCTTG CCCCCAAGA AGAAACTGCC 420  
 70 TTCGCTCAAG CAGGTGAACT CTGCCAGGAA GCAGCTGAGG CCCAAGGCCA CCTCCGAGC 480  
 CACTGTCCAA AGGGCAGGGT CCCAGCCAGC GTCCCAGGGC CTAGATCTCC TCTCCTCCTC 540  
 CACGGAGAAG CCTGGCCAC CGGGGGACCC GGACCCCATC GTGGCCTCCG AGGAGGCATC 600  
 AGAAGTGCCC CTTTGGCTGG ACCGAAAGGA GAGTGCCTGC CCTACAACAC CCGCACCCCT 660  
 GCAAATCTCC CCCTTCACTT CGCAGCCCTA TGTGGCCAC ACACTCCCCC AGAGGCCAGA 720  
 75 ACCCGGGGAG CCTGGGCCCTG ACATGGCCCA GGAGGCCCCC CAGGAGGACA CCAGCCCAT 780  
 GGCCCTGATG GACAAAGGTG AGAATGAGCT GACTGGGTCA GCCTCAGAGG AGAGCCAGGA 840  
 GACCACTACC TCCACCATTA TCACCACCAG GGTCATCACC ACCGAGCAGG CACCAGCTCT 900  
 CTGCAGTGTG AGCTTCTCCA ATCCTGAGGG GTACATTGAC TCCAGCGACT ACCCACTGCT 960  
 GCCCCTCAAC AACTTTCTGG AGTGACACATA CAACGTGACA GTCTACACTG GCTATGGGGT 1020  
 80 GGAGCTCCAG GTGAAGAGTG TGAACCTGTC CGATGGGGAA CTGCTCTCCA TCCGCGGGGT 1080  
 GGACGGCCCT ACCCTGACCG TCCTGGCCAA CCAGACACTC CTGGTGGAGG GGCAGGTAAT 1140  
 CCGAAGCCCC ACCAACACCA TCTCCGTCTA CTTCCGGACC TTCCAGGACG ACGGCCTTGG 1200  
 GACCTTCCAG CTTCCTTACC AGGCCTTCAT GCTGAGCTGC AACTTTCCCC GCCGGCCTGA 1260  
 CTCTGGGGAT GTCACGGTGA TGGACCTGCA CTCAGTGGGG GTGGCCCACT TTCCTGCCA 1320  
 85 CCTGGGCTAT GAGCTCCAGG GCGCTAAGAT GCTGACATGC ATCAATGCCT CCAAGCCGCA 1380  
 CTGGAGCAGC CAGGAGCCCA TCTGCTCAGC TCCTTGTGGA GGGGCAGTGC ACAATGCCAC 1440  
 CATCGGCCGC GTCCTCTCCC CAAGTTACCC TGAAAACACA AATGGGAGCC AATTCTGCAT 1500  
 CTGGACGATT GAAGTCTCCAG AGGGCCAGAA GCTGCACCTG CACTTTGAGA GGCTGTTGCT 1560



5  
10  
15  
20  
25

GCATGACAAG GACAGGATGA CGGTTACACAG CGGGCAGACC AACAAAGTCAG CTCTTCTCTA 1620  
CGACTCCCTT CAAACCGAGA GTGTCCCTTT TGAGGGCCTG CTGAGCGAAG GCAACACCAT 1680  
CCGCATCGAG TTCACGTCCG ACCAGGCCCG GGCGGCCTCC ACCTTCAACA TCCGATTTGA 1740  
AGCGTTTGTAG AAAGGCCACT GCTATGAGCC CTACATCCAG AATGGGAACT TCACTACATC 1800  
CGACCCGACC TATAACATTG GGACTATAGT GGAGTTCACC TGCAGCCCCG GCCACTCCCT 1860  
GGAGCAGGGC CCGGCCATCA TCGAATGCAT CAATGTGCGG GACCCATACT GGAATGACAC 1920  
AGAGCCCCTG TGCAGAGCCA TGTGTGGTGG GGAGCTCTCT GCTGTGGCTG GGGTGGTATT 1980  
GTCCCCAAC TGGCCCGAGC CCTACGTGGA AGGTGAAGAT TGTATCTGGA AGATCCACGT 2040  
GGGAGAAGAG AAACGGATCT TCTTAGATAT CCAGTTCCTG AATCTGAGCA ACAGTGACAT 2100  
CTTGACCATC TACGATGGCG ACGAGGTCAT GCCCCACATC TTGGGGCAGT ACCTTGGGAA 2160  
CAGTGGCCCC CAGAACTGT ACTCCTCCAC GCCAGACTTA ACCATCCAGT TCCATTCCGA 2220  
CCCTGCTGGC CTCATCTTTG GAAAGGGCCA GGGATTTATC ATGAACTACA TAGAGGTATC 2280  
AAGGAATGAC TCCTGCTCGG ATTTACCCGA GATCCAGAAT GGCTGGAAAA CCACTTCTCA 2340  
CACGGAGTTG GTGCGGGGAG CCAGAATCAC CTACCAGTGT GACCCCGGCT ATGACATCGT 2400  
GGGGAGTGAC ACCCTCACCT GCCAGTGGGA CCTCAGCTGG AGCAGCGACC CCCCATTTTG 2460  
TGAGAAAATT ATGTACTGCA CCGACCCCGG AGAGGTGGAT CACTCGACCC GCTTAATTTT 2520  
GGATCCTGTG CTGCTGGTGG GGACCAACAT CCAATACACC TGCAACCCCG GTTTTGTGCT 2580  
TGAAGGGAGT TCTCTTCTGA CCTGCTACAG CCGTGAAACA GGGACTCCCA TCTGGACGTC 2640  
TCGCCTGCC CACTGCGTTT CAGAAGCGGC AGCAGAGACG TCGCTGGAAG GGGGGAACAT 2700  
GGCCCTGGCT ATCTTCATCC CGTCCCTCAT CATCTCCTTA CTGCTGGGAG GAGCCTACAT 2760  
TTACATCACA AGATGTCGCT ACTATTCCAA CCTCCGCTG CCTCTGATGT ACTCCCACCC 2820  
CTACAGCCAG ATCACCCTGG AAACCGAGTT TGACAACCCC ATTTACGAGA CAGGGGGAAC 2880  
CCAAAAGGTT TAGGGTTTCA TTAAAAAGA GGTACCCTTT AAAAAGGGGC TTGTGAACTC 2940  
AACCCCAATT TCCCCGAGAC ATTTATCCAA AGGCCCTGGG GGCCTTGATT TAAACCCCA 3000  
AAAGGCGGCT GTTTTTTGGT TAAACTTTT AACAAAGGT TACGGGTTT TTCCCCGGAT 3060  
TTTATAAATT TTTAAAGTG

30 Seq ID NO: 208 Protein sequence:  
Protein Accession #: NP\_066938

35  
40  
45

1 11 21 31 41 51  
| | | | |  
MAQEAPQEDT SPMALMDKGE NELTGSASEE SQETTTSTII TTTVITTEQA PALCSVSFSN 60  
PEGYIDSSDY PLLPLNNFLE CTYNVTVTYTG YGVELQVKS V NLSDGELLSI RGVDGPTLTV 120  
LANQTLLEVEG QVIRSPNTNI SVYFRFTQDD GLGTFQLHYQ AFMLSCNFPR RPDSDGVTVM 180  
DLHSGGVAHF HCHLGYELQG AKMLTCINAS KPHWSSQEP I CSAPCGGAVH NATIGRVLSP 240  
SYPENTNGSQ FCIWTIEAPE GQKLHLHFER LLLHDKDRMT VHSQGTNKSA LLYDSLQTES 300  
VPFEGLLSEG NTIRIEFTSD QARAASTFNI RFEAFKGC HC YEPYIQNGNF TTSDDPTYNIG 360  
TIVEFTCDPG HSLEQGPAIL ECINVRDPYV NDTEPLCRAM CGGELSAVAG VVLSPNWPEP 420  
YVEGEDCIWK IHVGEEKRIF LDIQFLNLSN SDILTIYDGD EVMPHILGQY LGNSGPQKLY 480  
SSTPDLTIQF HSDPAGLIFG KGQGFIMNYI EVSRNDSCSD LPEIQNGWKT TSHTELVRGA 540  
RITYQCDPGY DIVGSDTLTC QWDLWSWSDP PFCEKIMYCT DPGEVDHSTR LISDPVLLVG 600  
TTIQYTCNPG FVLEGSSLLT CYSRETGTPI WTSRLPHCVS EAAETSLEG GNMALAIPIP 660  
VLIISLLLG AYYIYITRCRY YSNRLPLMY SHPYSQITVE TEFDNPIYET GGTQKV

50 Seq ID NO: 209 DNA sequence  
Nucleic Acid Accession #: NM\_001327.1  
Coding sequence: 89-631

55  
60  
65

1 11 21 31 41 51  
| | | | |  
AGCAGGGGGC GCTGTGTGTA CCGAGAATAC GAGAATACCT CGTGGGCCCT GACCTTCTCT 60  
CTGAGAGCCG GGCAGAGGCT CCGAGGCCAT GCAGGCCGAA GGCCGGGGCA CAGGGGGTTC 120  
GACGGGCGAT GCTGATGGCC CAGGAGGCCC TGGCATTCCT GATGGCCAG GGGGCAATGC 180  
TGGCGGCCCA GGAGAGGCGG GTGCCACGGG CGGCAGAGGT CCCCAGGGCG CAGGGGCAGC 240  
AAGGGCCTCG GGGCCGGGAG GAGGCGCCCC GCGGGGTCCG CATGGCGGCG CGGCTTCAGG 300  
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGGAG AGCCGCCTGC TTGAGTTCTA 360  
CCTCGCCATG CCTTTCGCGA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420  
GGATGCCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACCTG TGTCGGGCAA 480  
CATACTGACT ATCCGACTGA CTGCTGCAGA CCACCGCCAA CTGCAGCTCT CCATCAGCTC 540  
CTGTCTCCAG CAGCTTTCCC TGTGTATGTG GATCACGCGAG TGCTTTCTGC CCGTGTTTT 600  
GGCTCAGCCT CCCTCAGGGC AGAGGCGCTA AGCCAGCCTT GGCGCCCTT CCTAGGTCAT 660  
GCCTCCTCCC CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720  
GTTTGTGCTT GGAGGAGGAC GGCTTACATG TTTGTTTCTG TAGAAAATAA AACTGAGCTA

70 Seq ID NO: 210 Protein sequence:  
Protein Accession #: NP\_001318.1

75

1 11 21 31 41 51  
| | | | |  
MQAEGRGTTG STGDADGPGG PGIPDGPNGN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60  
PRGPHGGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EAELARRSLA QDAPPLPVP 120  
VLLKEFTVSG NILTIRLTAA DHRQLQLSIS SCLQLSLLM WITQCFLPVF LAQPPSGQRR

80 Seq ID NO: 211 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 52-459

85

1 11 21 31 41 51  
| | | | |  
CCTCGTGGGC CCTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGGAGC CATGCAGGCC 60  
GAAGGCCAGG GCACAGGGGG TTCGACGGGC GATGCTGATG GCCCAGGAGG CCCTGGCATT 120  
CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CGGGTGCCAC GGGCGGCAGA 180  
GGTCCCCGGG GCGCAGGGGC AGCAAGGGCC TCGGGGCCGA GAGGAGGCGC CCCGCGGGGT 240  
CCGCATGGCG GTGCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTGCGGGGC CAGGAGGCCG 300



5 GACAGCCGCC TGCTTCAGTT CCGACTGACT GCTGCAGACC ACCGCCAACT GCAGCTCTCC 360  
ATCAGCTCCT GTCTCCAGCA GCTTTCCTTG TTGATGTGGA TCACGCAGTG CTTTCTGCCC 420  
GTGTTTTTGG CTCAGGCTCC CTCAGGGCAG AGGCGCTAAG CCCAGCCTGG CGCCCCCTCC 480  
TAGGTCATGC CTCCTCCCCT AGGGAATGGT CCCAGCACGA GTGGCCAGTT CATTGTGGGG 540  
GCCTGATTGT TTGTCGCTGG AGGAGGACGG CTTACATGTT TGTTCCTGTA GAAAATAAAG 600  
CTGAGCTA

10 Seq ID NO: 212 Protein sequence:  
Protein Accession #: Eos sequence

15 1 11 21 31 41 51  
| | | | | |  
MQAEGQGTGG STGDADGPGG PGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60  
PRGPHGGAAS AQDGRCPGCA RRPDSRLLQF RLTAADHRQL QLSISSCLQQ LSLLMWITQC 120  
FLEPVFLAQAP SGQRR

20 Seq ID NO: 213 DNA sequence  
Nucleic Acid Accession #: NM\_000555  
Coding sequence: 416..1498

25 1 11 21 31 41 51  
| | | | | |  
CTTATTTTTT ATGAATGTCG GATAGCTGCA CCAGCTTGCT GGGGAAAGGG TTGATGAAT 60  
AGCACAAAGA CACTGGCTGT TCCCTGGAGG CTGTCCCTTT AAAGGAGAAT CTTAGTTTAT 120  
TCTGGGGGGA GGGGATGCAC ACATTAGAGT AGGAAAGAGG GCTTGGAATA AAATGAAAAC 180  
ACTCCCCCTT CATAGTCATT GTACTGAAAT GCAAAGACTG CTTCTTAAGC TGGAGATGCT 240  
AACCTTGGGT AGCTCCTTCT GTTCTCTTCA AGGGGAATTT TGTCAAGGCTA TGGATTCATT 300  
TACAACTGTT AGTCATGTGG GCATGTGTGA GGAAACAGAT GCCAGTTTTA ATGTATTTAG 360  
CCCGAAGTTC CAATTTGATA GGAGCCACTG TCAGTCTCTG AGGTTCCACC AAAATATGGA 420  
ACTTGATTTT GGACACTTTG ACGAAAGAGA TAAGACATCC AGGAACATGC GAGGCTCCCG 480  
GATGAATGGG TTGCCTAGCC CCACTCACAG CGCCCACTGT AGCTTCTACC GAACCAGAAC 540  
CTTGCAAGCA CTGAGTAATG AGAAGAAAGC CAAGAAGGTA CGTTTCTACC GCAATGGGGA 600  
CCGCTACTTC AAGGGGATTG TGTACGCTGT GTCTCTGAC CGTTTTTCGA GCTTTGACGC 660  
CTTGCTGGCT GACCTGACGC GATCTCTGTC TGACAACATC AACCTGCCTC AGGGAGTGCG 720  
TTACATTTAC ACCATTGATG GATCCAGGAA GATCGGAAGC ATGGATGAAC TGGAGGAAGG 780  
GGAAAGCTAT GTCTGTTTCT CAGACAACCT CTTTAAAAAG GTGGAGTACA CCAAGAATGT 840  
CAATCCCAAC TGGTCTGTCA ACGTAAAAAC ATCTGCCAAT ATGAAAGCCC CCCAGTCCTT 900  
GGCTAGCAGC AACAGTGCAC AGGCCAGGGA GAACAAGGAC TTTGTGCGCC CCAAGCTGGT 960  
TACCATCATC CGCAGTGGGG TGAAGCTGTG GAAGGCTGTG CGTGTGCTTC TGAACAAGAA 1020  
GACAGCCAC TCTTTTGAGC AAGTCTCTAC TGATATCACA GAAGCCATCA AACTGGAGAC 1080  
CGGGGTGTG AAAAACTCT AACTCTGGA TGGAACACAG GTAACCTGTC TCCATGATTT 1140  
CTTTGGTGAT GATGATGTGT TTATTGCCTG TGGTCTGAA AAATTTTCGT ATGCTCAGGA 1200  
TGATTTTTCT CTGGATGAAA ATGAATGCCG AGTCATGAAG GGAAACCCAT CAGCCACAGC 1260  
TGGCCCAAG GCATCCCCAA CACCTCAGAA GACTTCAGCC AAGAGCCCTG GTCCTATGCG 1320  
CCGAAGCAAG TCTCCAGCTG ACTCAGCAA CGGAACCTCC AGCAGCCAGC TCTCTACCCC 1380  
CAAGTCTAAG CAGTCTCCCA TCTCTACGCC CACCAGCTCT GGCAGCCTCC GGAAGCACAA 1440  
GGACCTGTAC CTGCCTCTGT CTTGGATGA CTCGACTCG CTTGGTGATT CCATGTAAAG 1500  
GAGGGGAGAG TGCTCAGAGT CCAGAGTACA AATCCAAGCC TATCATTGTA GTAGGGTACT 1560  
TCTGCTCAAG TGTCCAACAG GGCTATTGGT GCTTTCAAGT TTTTATTTTG TTGTTGTTGT 1620  
50 TATTTTGAAA AACACATTGT AATATGTTGG GTTTATTTTC CTGTGATTTT TCTCTGGGC 1680  
CACTGATCCA CAGTTACCAA TTATGAGAGA TAGATTGATA ACCATCCTTT GGGGCAGCAT 1740  
TCCAGGGATG CAAAATGTGC TAGTCCATGA CTTTCAATG GAAAGCTTAG GGCCTGGGG 1800  
TAAATTTGCC CCGTTTAAAT TTGCCCCAAC AGTTTTCTCT TTGTAGAGGG GTGTTTAAAT 1860  
55 ATACAGCAAT TAAAAAGTTT GTGTGGGGAA AAAAAAACT CATTGGCAGA TCCAAGAATG 1920  
ACAAACACAA GTGCCCCTTT TCTCTGGATC TCAAGAATGG TGGAGGACCC TGAAGGACA 1980  
GCAAGGCAGC TCCCCAGCCT CACTCTTCAC TCCTGATTGA GGCCCGGGTT TGTGTGCCAG 2040  
CACCAATTCT GGCTGTCAAT GGGGAGAAAT AAACCAACAA CTTATAATTG TGACACCAGA 2100  
TGCTTAGGAT CTTGGTGTG GGTAGCTTAA GAGAATAGAC AGAATTGGAA AATACTGCAG 2160  
60 ACATTTCCGA AGAGTTTATA AAGCACAGTG AATTCTCTGG CAATCTCTCC ACTGAGGCAA 2220  
TTTGGAATCA ATAAGCAATT GATAATAGTT TGGAGTAAGG GACTTCATAT ACCTGATTCC 2280  
TCTAGAAGGC TGTCTAACAT ACCACATGAT TACATGAAC GTATGGTATC CATCTATCTC 2340  
TGTTCTATTG AATGCCTTGT TAACAGCCAA CACTGAAAAC ACTGTGAGAA TTTGTTTTCA 2400  
GGTCTGACAC CTTTTCAGTCT CTTTTTATAG CAAGAAATCA ATATCCTTTT TATAAAAATT 2460  
65 CATGCTGTGA TTTTCAGGAG AAACCTCTTCA GGCTCCTTTT TTATAAACTG GTGATTTTTC 2520  
TTTTGTCTAA AAAACACATG AAGAAAATTT ACCAGAAAAA AAAAAAAAAG CCGAAGAATA 2580  
ATGTTATTTA GAAATTATGC TGTCAGTGCC AAACAGTAAC CTCCAGGAGA AAACAAGATG 2640  
AATAGCAGAG GCCAATTCAA TAGAATCAGT TTTTGTATAG CTTTTTAACA GTTATGCTTG 2700  
CATTAATAAT TTCAATGTGG ACCAGACATT CTAATTATAT TTTAAATGAA ATGTTACAGC 2760  
70 ATATTTTAAG CAACTCTTTT TATCTATAAT CCTAATATTT CATACTGAAG ACACAGAAAT 2820  
CTTTCACCTG TCTTTAACAT TAGAAAGGAT TTCTCTTTAC TAAGGACTGA TCATTTGAAA 2880  
TAGTTTTTCA TCTTTTGAGA TACAGGTTTA TAACACTGCT TTTTTTTTCC TGTAAACATA 2940  
GCCATAATG GCAAAAACAA CTAATTTTAA TTGAAGGTCT TGCTTGCCAN TCCTGTGTTG 3000  
GCTTTNACCA AATATAAAAA TTCCCTTAT CTTTGGTAAT GGTGCAATN TTTGGAAAGG 3060  
75 CACAGCATCC AAACCAAGCT GCTGTTTGGC TACTGAATGG CTTGCAGTTG TTCTCCACT 3120  
CTAAATGGAA TGAGCTTGCT GTGTGTGTGT GTGGTGGTGG TGGGAGGGGG TGGTGATGTT 3180  
GTGTGTGTGT GTGTGCATCT GCAGCTGCTT CAAAATTAAG AAATACTACA AGACACCCCT 3240  
GTAATGGATT GGTGGCAACT GGTGGCACT CTGTATGTGC ACTGTGTAGG GGGGAACCCA 3300  
GTGGTGGTGG GGTATCTCAA ATGCCCTAG ACAAGCTTCA GATGTCTGTA GCTACCAAAA 3360  
80 ACATTTTCGG TTCAAGAAAA GTGAGATGAT GGTAGTACTG GTTCTGGTGT AAATTGAAAA 3420  
ACCCCAAAAT ATGAGGATCT CTTTTTGCCC CCTCTCCTTT TTTTGTAAAC CCATTCAAAA 3480  
CCATTAATAA GCCCATTTTA CTAANCCCTT ATTTCTTTCT AGAAGCTCAG GGTTCNCTTA 3540  
GTGCCTCCCA NAACATTTTG TAGTTAATTG GGAATAAGTG ATACTTGGAT TAGGGGGTGT 3600  
GGCATAAAG AATGGTGGGA GGCCTGATTG TAAAATTCAG GCCAGAACCC CCAATGACTC 3660  
75 CACCCATAGT NTCATTTTAG GTCTCATTTA GTCCATCACC TTTATTTTAA GTTGAGGAAG 3720  
85 TGGAGGCTGG TAAAGAGCAG GACCAGAGGA AGAATCCAGA TTTCTTTATG CTTGGGCCTC 3780  
ACACTAGCTC TGTGAGTATT TCCTTGATTG CCGTATATGT ACTACTAGAA AATACCAAAT 3840  
GGATATATTT TCTTTAGGAT AACCTTTGAA CCAACAATNT TCAATAACAA TAGTACATCT 3900



	TCCATCTTAC	TTTTAATCGA	GTATAAGGAA	ATGTTTCTTT	ATGGCCATTT	TGGAGGGAGC	3960
	AGGGGATGAG	GCTTGGCATA	GTCCAAAATT	TAAGNCTCCA	ATAATTAATT	GCATTTTAAA	4020
	TTGTTTTTAAA	TTGGCCCACT	TTCAAGGCAG	TTTTTTTTGT	GTGTCTGTAA	CTGAGCTCCT	4080
5	CCACCCCTGT	CATTCACTTC	CAATTTTACC	CAATCCAATT	TTAGCACTCA	AGTTCCATTG	4140
	TGTTAATTTT	TGCACGGTCT	ACACACATCA	AGTCAGCAAG	CATTTGCCAC	CACTCCCTAT	4200
	ACTTCTCCCT	CTTTTTTACA	CACACACACA	CACACACACA	CACAATCCAT	CTCTTGCTTG	4260
	TTCTACCTC	CCTGATTTTT	CTTCCCTACA	GAAATAGAAA	TAGGGACAAA	GAAGGGGAAA	4320
	ATGTATATAT	TGGGGCTGGG	CTGAACAACT	AACTTCATAA	GTAGTATTAA	CTAGGGGTAA	4380
10	ATTGAGAGAA	AAGCTCCTTT	TCTCTTCACT	GTTTTGGAAA	GGATAGCCAT	TAGCATGACT	4440
	GCTTTGTGTC	CTTATGGACT	TTAGTATTAG	CCTAGATTGA	ATTATAGCGT	TTTTCTAGCT	4500
	GAAGGAACCT	TAAGATCACA	TCATCTACTC	CTCTACTCCA	AATTTCTCAT	TCTTCAGGCC	4560
	AGGAAACCGA	GACACAGAGG	TAAAGTAATT	TCCCCAAGGT	CACACAGCTG	GCTGGGGCAG	4620
	GATTGGGTTT	ACAACCCACA	TCTCCTGGCT	CTTATTCCAG	GGCCTTTTCC	CACTAAGTAG	4680
15	TATTGCCTTC	CATTAGGCTC	CTGAGAGTTA	TTTCTCAGGG	TCATGTTGCA	TCTTGGAGCC	4740
	ACATGCTGCT	GCCCTGATCT	CAGTGGGAAA	TNCAACCCAGC	AACCTAATAC	AGCCCCTTTT	4800
	CCCTGCATTG	ACCTGGTTCC	CATCCACATG	GGTTGCAGAT	GTCTTGAAG	AGAGTGAGGC	4860
	ATTGAGGGCC	AATAGGAGCA	ATGGGGTCCC	TGGCCTTGTC	CATCTGATTC	AGGAGATCAC	4920
	TGCTCCATCG	TGAGGAGCCC	TCTGAATAGC	CCCCCACTGA	ATGCTTGCCCT	TGCCCCAATG	4980
20	GAATGGAGGA	AGATTGATTT	TCTCCATCAG	TTACCTTGT	GTCTCTCAT	AATGGTTGGT	5040
	CTTTCCAGGC	TGAGGGAAT	GTTTCTTGTT	TCCANAGTAN	AAAAAGAAA	GAGTGGAACA	5100
	ATANCTTTGT	TTCTCCTAAC	TTTCTGAGAT	GGCTTTTCAA	CATTTAAAAA	AACTAGTGT	5160
	GGTACCATTG	ACTGGCANGA	TTTNTTTAG	AATATTGGAG	TAAGATGAGG	TAGAGAAAAT	5220
	AACCTGGTCT	CACGTGTGTT	GCCCTCATCC	ACAATGTCCC	CAAAGCCATC	CTGCTNTGAT	5280
25	GAGGACAATT	TCCAGGTATA	AGCAAGGGGC	TTTGTGACAA	AAATGTACCC	TGGCTGATGT	5340
	TAAACATTGG	CTCCTGTGTT	TGCACCAAAA	TAGCAAGCTG	TGTGCTCTAT	ACACTCTTCC	5400
	CATCGTCTTG	TGTACACTGC	TCCTGTGGCC	TTCCACAGCA	GAAACCAGGG	CAAAAGGGTC	5460
	CAACACATG	GTTTTCCCTG	CTGCAAGGCT	NTTCTGGGA	ACTAAGGGGG	TATTTATTAG	5520
	TTCAGTTNTA	AGAGACCTCC	TTCTGGGCTT	ACCCCACTCC	TCAGGTACTT	CTCTCTCCTT	5580
30	CCTCCTTCTC	CTCCACAGTC	ACAAGTAACC	AAGGAACCTG	AAAGTGGATG	TGTAGCTATT	5640
	TGAAGAAGGC	AAGGAACCCCT	GAGATTCTTC	TTTGAATCCT	TTAGTCCAAG	TCTTAGACCA	5700
	GTGATTGGTG	CTTACCTTGA	ACAAAATTTT	GTCTGTGTTT	CTAATCCCTT	CAATACTNTG	5760
	GGTACAATGC	TCCCAATCAC	CCTGCACATT	TGATTCTAAA	TGGCTTTTAT	TTTTTAAAAA	5820
	TCCATATCCC	TAGGACAAGA	NAACAGGATG	CCTATATCCC	CAAAATGAGC	TCCAGGACAC	5880
35	TGATGGGAAT	GATCCCAANG	ATCACCACAC	CTCAGAAAAC	GTCTGTGCCA	ANAGACTTCC	5940
	CCAGATAGAA	NCACTGGGAC	AGTGGTTTGA	ACGACTTCTT	TTATGGTTGT	CCAGTTTGCT	6000
	ATGGAAATAA	AAGGCATTGA	TTTTTTAAAA	AAGATGATTG	GAACCTGTCT	TTGGCCACAT	6060
	AGGGCCACTT	GGATCCATTT	CCAGGCCTTA	CTCATATATT	GCCTTCACTG	AAGGGCTTTG	6120
	GCTTTAAGTC	CCAGACTGGT	CTCCCAAGTG	AACCATAAGT	GTTTTGGAGC	TCATCTGGGG	6180
40	TGAGGCATGA	GAATGTTGCC	CCATCTATCT	CTTCAGGAAA	AGGTGCCCTC	CCTCCCTTTC	6240
	TCCTAAAGCC	TGGTCCCAAA	AAATGTGTTT	TGCTCCCAA	AGTCTAGTAT	GGTCTTTATA	6300
	CACCCANACT	CTTAGTGTG	CGTCTGCCT	TGTTTCTTGG	TTAAGGATCT	ATGCANACCT	6360
	CCCGCTTTGG	CTTAGCTAGC	GTGACATTGG	CTATCATTGG	ACAAGACTAA	CTTTTTTTTT	6420
	TTTTTTTTTTG	ACTGAGTCTC	CCTCTGTAC	CTAGGCTGGA	GTGCAGTGGC	ACAATCTTGG	6480
45	CTCGCTGCAA	CCTTCACCCCT	TCACCTCCCA	GGTCGAAGCG	ATTCTCCTGC	CTCAGTCTCC	6540
	CGAGTAGCTG	GGATTACAGG	CGTGCGCCAC	CAAATCTGGC	TATTTTTTTA	TTATTATTAT	6600
	TTTTAGTAGA	GATGGGTTT	CACCATGTTG	GCCAGACTGG	TCTTGAACCT	TTGGCCTCAA	6660
	ATTATCTGCC	CACCTCGGCC	TCCCAAGTG	CTGGGATTAC	AGGCATGAGC	ACCATGCCCA	6720
	GCTGACAAGA	CTAATTTTTT	ATCCCTTGGT	TTATTGGCTT	CAACATCTTC	TGGAATCAGA	6780
50	GGTGATTTTT	TCTTACCTTG	GATGCCCTGAG	ACTAGGGGAG	TATAGAAATC	CAATTGGTAA	6840
	TTAAGGCATC	TTTCTGCTCC	TGATCAGAAG	GGCAGGTTAG	TTGGGAGAGG	TCAGATGGCA	6900
	CAACAGAAGT	CACCTTGTA	GTAAGGCAAA	GACTTTGAAG	GCATTAGCGT	TTCTCATTAC	6960
	TTAGGTCAAT	AACCTTGAGG	GAATCAATGG	CTTTTTTGCC	GCTCTACCTC	TTTGTGTATC	7020
	TCTTTGACTT	TTCTTTCTCT	GTCTGTTTCT	CTCTGTTCTC	AGTTTATATT	CTATGTTATC	7080
55	AGTCTCTCTT	TCCACAGTAC	AAACATCCAT	CCTTTCTCCT	GTGCAATTCT	GTCTCTCCCT	7140
	CTTATTATCT	TTATTTGTAC	TTTTTCCTTC	CTCCCTGTCT	AGGCATTGGG	CATGTGCCTC	7200
	TTCTTAGCCT	GTGATTTTGC	CTTGGGACTG	ATGATAAATT	ATTTCCAGAT	TCAATCAGCC	7260
	CTGGTCCTAC	CCCAGTCCAA	TCAGAAGTAT	GTTGGTGGGG	AATCAACCTG	ATCCTGGCCC	7320
	TTTCTTCTTC	TCCATTTTCA	TTCTGTAATC	CCCTCAGCAG	ATCTTTACAA	GCAGTTTCCT	7380
60	TCTAGTCAT	GTATCTTTAG	GTCTTTGCCT	TCCAAGCACT	GTACAGAATA	CTTTGTGGTT	7440
	CCTTTTTTAGT	CTGACATTTT	GTGGAGCAGT	GAAGCGTGCT	CAGAGACATA	ATCAGCTGAA	7500
	GAGAAAAAAT	CCACCCATGG	ATTTATATCA	GCTAAATACT	AATAATTGAT	TTTGTTTGAT	7560
	GTGCCCATAA	TTTTTAAAGC	TGCAATATAA	TATAATGAGG	GACCACAGGT	AATTTCTCCT	7620
	GTCATTTGTT	TTGGCTGGAT	GGGGGTGGGG	GAGTAATTGC	TTAAAGTTTT	ACCATTACAC	7680
65	ATTAAACTCT	CTATAATAAT	CTTGTTTGGG	GCTTGCTAAC	TGTTGAGCTG	TTTTAACTAA	7740
	ACTGGTAGGC	AATCGGAGTT	GATTTAAATG	AAAAGATAAT	TTAACAAATC	TATACTATAA	7800
	AAAGAGACAT	TTGCTTAATT	GACATGTATT	TTTTCTTCT	GAGTCACCTA	AACATTTACT	7860
	CTTGACACCA	ACTGTTCATG	ATACTGAATA	GACAGTCCAT	ATAAGAGAAA	TTAGTGGACC	7920
	TAAAGAAGCC	AGATTGTAGG	TGTTAATTTA	TTAAACAGAA	TTGCAAAGCC	CTTGGAATG	7980
70	TCACTGCTTG	GCAATACCAT	ATGGCATGCC	AAAATTTACA	ATGACTTTTC	TTTATAAGTT	8040
	ATCCAAAAGG	GATTTGAACA	AGTAAGAGGT	TATGCCAAAA	TGTCTCCAAT	GTATGGTCCT	8100
	GTAATATATT	GCAGCTTGAA	GCCAATGATC	CCTTATGACT	TGTATACAAC	TAATGCATGT	8160
	TTTATTGAAT	TTTGCAATTC	CCACGTGTGG	TAAGTCTTTA	AAATGTTTTT	GATCACCTTT	8220
	NTGTGCCATT	AACTTGTAC	AGAAAAATGT	TTTATGGCCA	TTTTCAAAGG	GAGAAAGTTT	8280
75	AAAATGGAAA	CAGCCCACCC	TTTCTGCCCT	ATAGCTGTAG	TTAGAATTGA	GTACCTGTAG	8340
	CAAAACAGCT	GTAATTGGTG	GTTGTAGTGT	TAGAGGTGTT	AGCTTGCTAG	TGACTAGCTT	8400
	TGGAGAGTAA	ATGCATGGTA	TTGTACATCA	CATTTCTTAA	CTCGTTTTAA	CCTCTGAAAA	8460
	GAATATATTG	TTCTTTGTAG	TCCTTCTTCC	CACCCCTTGG	CCCTCTCCCT	CTCCCTGCTC	8520
	CCAGTTGTCT	TACAGTTGTA	AATATCTGAT	TTGAGGCCCA	ATAACTCTTG	CCAAGTAAAG	8580
80	TCAGCAAACA	ACAAACAAAC	CAAAATGTGG	GGAAAAGGCA	TTTCTCAACC	ATCTCTCAGC	8640
	AGTTATTGAT	CATTTCTTAA	GGAACAGCAT	TGTGATCAAA	GACTCAACTT	TACGTAAAAA	8700
	TCAGTGGTAA	ATTGGGGTTG	TATTGGCCAT	TGATTACATT	CAGGATTGAA	TAGTTTTTCA	8760
	AATCAGATGT	AATCCAAAGA	CAGTAGGTAG	TGATGTCCCT	TATCCCTGCA	GCTGTTTTAA	8820
	GATAGAGACC	TCAGAAGACT	CTGCTTGACC	GATGACCAAT	AATTATTGTA	AAAAAAAAGA	8880
85	AAAAATGAGA	GAAATAAAAC	AGATATTTAA	GAACTTTAGC	CACCTATTTA	GAATAGTTAT	8940
	AGCCAGAAAA	AAAAACAAGG	GCAATGAGTT	AAATGCATTA	CTATCAGTGT	CCTAGGCAAT	9000
	ACCTAACCTA	CTCTGAAATT	GTGATTCAAA	AGCAGTATTT	CAAGAGGCAT	TCTCCTTTTT	9060
	TGGTTTGCTG	ACCCCACTTG	GACTGGTAGG	TTTGGTGAGG	CCCCCATAAA	CCAGCTGGAG	9120



5 CAGACCCCTT TCATCTCCTG TGCCTGTAAC ACCCCTCTTC CCCACCCCC TCCGCAATTC 9180  
AATGAGGGCT TTCTTGGGTC AGAGGACTTC AAGGTTGTCT AGAGAAGTTT GCCATGTGTG 9240  
TAAGGTGCTG TGAACGTGTA GTGCTGAAGA TTCGAGCAT TCAATACCAG GCAGCCAAAG 9300  
AGCTGCTCTT GCAATTATT TGGCTCTCAA GCTCTGTTCT TCATCGCATT CTCATTTCTG 9360  
TGTACATTTG CAAGATGTGT GTAATGTCAT TTTCCAAAA TAAATTTGA TTTCAAT

10 Seq ID NO: 214 Protein sequence:  
Protein Accession #: NP\_000546

15 1 11 21 31 41 51  
MELDFGHFDE RDKTSRNMRG SRMNGLPSP T HSAHCSFYRT RTLQALSNEK KAKKVRFYRN 60  
GDRYFKGIVY AVSSDRFRSF DALLADLTRS LSDNINLPQG VRYIYTIDGS RKIGSMDELE 120  
EGESYVCSSD NFFKKVEYTK NVNPNWSVNV KTSANMKAPQ SLASSNSAQA RENKDFVRPK 180  
LVTIIRSGVK PRKAVRVLLN KKTAKSFEQV LTDITEAIKL ETGVVKKLYT LDGKQVTCLE 240  
DFFGDDDVFI ACGPEKFRYA QDDFSLDENE CRVMKGNPSA TAGPKASPTP QKTSAKSPGP 300  
MRRSKSPADS ANGTSSSQLS TPKSKQSPIS TPTSPGSLRK HKDLYLPLSL DDSDSLGDMS

25 Seq ID NO: 215 DNA sequence  
Nucleic Acid Accession #: NM\_130467  
Coding sequence: 312..644

30 1 11 21 31 41 51  
GGCACGAGGC AGAGCTCTGC AAGGAGAGGT TGTGTCTTCG TTCTTTCCGC CATCTTCGTT 60  
CTTTCCAACA TCTTCGTTCT TTCTCACTGA CCGAGACTCA GCCGGTAGGT CTGCAGAGTG 120  
GTCTTCCTGG TAATTTAGTT GTGAGTGAAT GTGTGGAGGA GCCAGCGGGC TTAGGACAGG 180  
TCCTGTGGCA CAGTCCGTGG CTTTGAGGGA AAAGGGCCTC GCGGTGGTCC TCCGCCTTCC 240  
CCCAGGTCGT GATGCAGGCG CCATGGGCGG GTAATCGTGG CTGGGCTGGA ACGAGGGAGG 300  
AAGTGAGAGA TATGAGTGAG CATGTAACAA GATCCCAATC CTCAGAAAGA GGAAATGACC 360  
AAGAGTCTTC CCAGCCAGTT GGACCTGTGA TTGTCCAGCA GCCCACTGAG GAAAAACGTC 420  
AAGAAGAGGA ACCACCAACT GATAATCAGG GTATTGCACC TAGTGGGGAG ATCAAAAATG 480  
AAGGAGCACC TGCTGTTCAA GGGACTGATG TGGAAGCTTT TCAACAGGAA CTGGCTCTGC 540  
TTAAGATAGA GGATGCACCT GGAGATGGTG CTGATGTCAG GGAGGGGACT CTGCCCCACTT 600  
TTGATCCAC TAAAGTGCTG GAAGCAGGTG AAGGGCAACT ATAGGTTTAA ACCAAGACAA 660  
ATGAAGACTG AAACCAAGAA TATGTGTTCT ATGCTGGAAA TTTGACTGCT AACATTCTCT 720  
TAATAAAGTT TTACAGTTTT CTGCAAAAAA AAAAAAAAAA AAA

45 Seq ID NO: 216 Protein sequence:  
Protein Accession #: NP\_569734

50 1 11 21 31 41 51  
MSEHVTRSQS SERGNDQESS QPVGPIVQQ PTEEKQREEE PPTDNQGIAP SGEIKNEGAP 60  
AVQGTLDVEAF QQELALLKIE DAPGDGPDVR EGTLPFTDPT KVLEAGEGQL

55 Seq ID NO: 217 DNA sequence  
Nucleic Acid Accession #: NM\_001476.1  
Coding sequence: 82..435

60 1 11 21 31 41 51  
GCCAGGGAGC TGTGAGGCAG TGCTGTGTGG TTCCTGCCGT CCGGACTCTT TTTCCTCTAC 60  
TGAGATTCAT CTGTGTGAAA TATGAGTTGG CGAGGAAGAT CGACCTATTA TTGGCCTAGA 120  
CCAAGGCGCT ATGTACAGCC TCCTGAAGTG ATTTGGGCTA TGCGGCCCGA GCAGTTTCAGT 180  
GATGAAGTGG AACCAGCAAC ACCTGAAGAA GGGGAACCAG CAACTCAACG TCAGGATCCT 240  
GCAGCTGCTC AGGAGGGAGA GGATGAGGGA GCATCTGCAG GTCAAGGGCC GAAGCCTGAA 300  
GCTGATAGCC AGGAACAGGG TCACCCACAG ACTGGGTGTG AGTGTGAAGA TGGTCTGAT 360  
GGGACGGAGG TGGACCCGCC AAATCCAGAG GAGGTGAAAA CGCCTGAAGA AGGTGAAAAG 420  
CAATCACAGT GTTAAAAGAA GACACGTTGA AATGATGCAG GCTGCTCCTA TGTTGGAAAT 480  
TTGTTTATTA AAATTCTCCC AATAAAGCTT TACAGCCTTC TGCAAAA

70 Seq ID NO: 218 Protein sequence:  
Protein Accession #: NP\_001467.1

75 1 11 21 31 41 51  
MSWRGRSTYY WPRPRRYVQP PEVIGPMRPE QFSDEVEPAT PEEGEPATQR QDPAAQEGE 60  
DEGASAGQGP KPEADSQEQG HPQTGCECED GPDGQEVDP NPEEVKTPEE GEKQSQC

80 Seq ID NO: 219 DNA sequence  
Nucleic Acid Accession #: NM\_001476  
Coding sequence: 90-3671

85 1 11 21 31 41 51  
ACAGCGGAGC GCAGAGTGAG AACCACCAAC CGAGGCGCCG GGCAGCGACC CCTGCAGCGG 60  
AGACAGAGAC TGAGCGGCCG GGCACCGCCA TGCCTGCGCT CTGGCTGGGC TGCTGCCTCT 120  
GCTTCTCGCT CCTCTGCCC GCAGCCCGGG CCACCTCCAG GAGGGAAGTC TGTGATTGCA 180  
ATGGGAAGTC CAGGCAGTGT ATCTTTGATC GGGAACTTCA CAGACAAACT GGTAATGGAT 240  
TCCGCTGCCT CAACTGCAAT GACAACACTG ATGGCATTCA CTGCGAGAAG TGCAAGAATG 300  
GCTTTTACCG GCACAGAGAA AGGGACCGCT GTTTGCCCTG CAATTGTAAC TCCAAAGGTT 360



	CTCTTAGTGC	TCGATGTGAC	AACTCTGGAC	GGTGCAGCTG	TAAACCAGGT	GTGACAGGAG	420
	CCAGATGCGA	CCGATGTCTG	CCAGGCTTCC	ACATGCTCAC	GGATGCGGGG	TGCACCCAAG	480
	ACCAGAGACT	GCTAGACTCC	AAGTGTGACT	GTGACCCAGC	TGGCATCGCA	GGGCCCTGTG	540
5	ACGCGGGCCG	CTGTGTCTGC	AAGCCAGCTG	TTACTGGAGA	ACGCTGTGAT	AGGTGTGCGAT	600
	CAGGTTACTA	TAATCTGGAT	GGGGGGAACC	CTGAGGGCTG	TACCCAGTGT	TTCTGCTATG	660
	GGCATTGAGC	CAGCTGCCGC	AGCTCTGCAG	AATACAGTGT	CCATAAGATC	ACCTCTACCT	720
	TTCATCAAGA	TGTTGATGGC	TGGAAGGCTG	TCCAACGAAA	TGGGTCTCCT	GCAAAGCTCC	780
	AATGGTCACA	GCGCCATCAA	GATGTGTTTA	GCTCAGCCCA	ACGACTAGAC	CCTGTCTATT	840
10	TTGTGGCTCC	TGCCAAATTT	CTTGGGAATC	AACAGGTGAG	CTATGGGCAA	AGCCTGTCCT	900
	TTGACTACCG	TGTGGACAGA	GGAGGCAGAC	ACCCATCTGC	CCATGATGTG	ATTCTGGAAG	960
	GTGCTGGTCT	ACGGATCACA	GCTCCCTTGA	TGCCACTTGG	CAAGACACTG	CCTTGTGGGC	1020
	TCACCAAGAC	TTACACATTG	AGGTTAAATG	AGCATCCAAG	CAATAATTGG	AGCCCCCAGC	1080
	TGAGTTACTT	TGAGTATCGA	AGGTTACTGC	GGAACTCTAC	AGCCCTCCGC	ATCCGAGCTA	1140
15	CATATGGAGA	ATACAGTACT	GGGTACATTG	ACAATGTGAC	CCTGATTTCA	GCCCGCCCTG	1200
	TCTCTGGAGC	CCCAGCACCC	TGGGTTGAAC	AGTGTATATG	TCCTGTTGGG	TACAAGGGGC	1260
	AATTCTGCCA	GGATTGTGCT	TCTGGCTACA	AGACGATATC	AGCGAGACTG	GGGCCTTTTG	1320
	GCACCTGTAT	TCCTTGTAAC	TGTCAAGGGG	GAGGGGCCCTG	TGATCCAGAC	ACAGGAGATT	1380
	GTTATTGAGG	GGATGAGAAT	CCTGACATTG	AGTGTGCTGA	CTGCCCAATT	GGTTTCTACA	1440
20	ACGATCCGCA	CGACCCCGGC	AGCTGCAAGC	CATGTCCCTG	TCATAACGGG	TTCAGCTGCT	1500
	CAGTGATGCC	GGAGACGGAG	GAGGTGGTGT	GCAATAACTG	CCCTCCCGGG	GTCCACGGTG	1560
	CCCGCTGTGA	GCTCTGTGCT	GATGGCTACT	TTGGGGACCC	CTTTGGTGAA	CATGGCCAG	1620
	TGAGGCCCTG	TCAGCCCTGT	CAATGCAACA	ATGGAGTGGT	CCCCAGTGCC	TCTGGGAATT	1680
	GTGACCCGCT	GACAGGCAGG	TGTTTGAAGT	GTATCCACAA	CACAGCCGGC	ATCTACTGCG	1740
25	ACCAGTGCAA	AGCAGGCTAC	TTCGGGGACC	CATTGGCTCC	CAACCCAGCA	GACAAGTGTC	1800
	GAGCTTGCAA	CTGTAACCCC	ATGGGCTCAG	AGCCTGTAGG	ATGTCGAAGT	GATGGCACCT	1860
	GTGTTTGCAA	GCCAGGATTT	GGTGGCCCCA	ACTGTGAGCA	TGGAGCATTC	AGCTGTCCAG	1920
	CTTGCTATAA	TCAAGTGAAG	ATTGAGATGG	ATCAGTTTAT	GCAGCAGCTT	CAGAGAATGG	1980
	AGGCCCTGAT	TTCAAAGGCT	CAGGGTGGTG	ATGGAGTAGT	ACCTGATACA	GAGCTGGAAG	2040
30	GCAGGATGCA	GCAGGCTGAG	CAGGCCCTTC	AGGACATTCT	GAGAGATGCC	CAGATTTTCA	2100
	AAGGTGCTAG	CAGATCCCTT	GGTCTCCAGT	TGGCCAAGGT	GAGGAGCCAA	GAGAACAGCT	2160
	ACCAGAGCCG	CCTGGATGAC	CTCAAGATGA	CTGTGGAAAG	AGTTCGGGCT	CTGGGAAGTC	2220
	AGTACCAGAA	CCGAGTTCGG	GATACTCACA	GGCTCATCAC	TCAGATGCAG	CTGAGCCTGG	2280
	CAGAAAGTGA	AGCTTCCTTG	GGAAACACTA	ACATTCCTGC	CTCAGACCAC	TACGTGGGGC	2340
35	CAATGGCTT	TAAAAGTCTG	GCTCAGGAGG	CCACAAGATT	AGCAGAAAGC	CACGTTGAGT	2400
	CAGCCAGTAA	CATGGAGCAA	CTGACAAAGG	AACTGAGGA	CTATTCCAAA	CAAGCCCTCT	2460
	CACGTGGTGC	CAAGGCCCTG	CATGAAGGAG	TCGGAAGCGG	AAGCGGTAGC	CCGGACGGTG	2520
	CTGTGGTGCA	AGGGCTTGTG	GAAAAATTGG	AGAAAACCAA	GTCCCTGGCC	CAGCAGTTGA	2580
	CAAGGGAGGC	CACTCAAGCG	GAAATTGAAG	CAGATAGGTC	TTATCAGCAC	AGTCTCCGCC	2640
40	TCCTGGATTG	AGTGTCTCGG	CTTCAGGGAG	TCAGTGATCA	GTCTTTTCAG	GTGGAAGAAG	2700
	CAAAGAGGAT	CAAAACAAAA	GCGGATTTCAC	TCTCAACGCT	GGTAACCAGG	CATATGGATG	2760
	AGTTCAAGCG	TACACAAAAG	AATCTGGGAA	ACTGGAAGGA	AGAAGCACAG	CAGCTCTTAC	2820
	AGAATGGAAA	AAGTGGGAGA	GAGAAATCAG	ATCAGCTGCT	TTCCCGTGCC	AATCTTGCTA	2880
	AAAGCAGAGC	ACAAGAAGCA	CTGAGTATGG	GCAATGCCAC	TTTTTATGAA	GTTGAGAGCA	2940
45	TCCTTAAAAA	CCTCAGAGAG	TTTGACCTGC	AGGTGGACAA	CAGAAAAGCA	GAAGCTGAAG	3000
	AAGCCATGAA	GAGACTCTCC	TACATCAGCC	AGAAGGTTTC	AGATGCCAGT	GACAAGACCC	3060
	AGCAAGCAGA	AAGAGCCCTG	GGGAGCGCTG	CTGCTGATGC	ACAGAGGGCA	AAGAATGGGG	3120
	CCGGGGAGGC	CCTGGAAATC	TCCAGTGAGA	TTGAACAGGA	GATTGGGAGT	CTGAACTTGG	3180
	AAGCCAAATG	GACAGCAGAT	GGAGCCTTGG	CCATGGAAAA	GGGACTGGCC	TCTCTGAAGA	3240
50	GTGAGATGAG	GGAAGTGGAA	GGAGAGCTGG	AAAGGAAGGA	GCTGGAGTTT	GACACGAATA	3300
	TGGATGCAGT	ACAGATGGTG	ATTACAGAAG	CCCAGAAGGT	TGATACCAGA	GCCAAGAACG	3360
	CTGGGGTTAC	AATCCAAGAC	ACACTCAACA	CATTAGACGG	CCTCCTGCAT	CTGATGGACC	3420
	AGCCTCTCAG	TGTAGATGAA	GAGGGGCTGG	TCTTACTGGA	GCAGAAGCTT	TCCCGAGCCA	3480
	AGACCCAGAT	CAACAGCCAA	CTGCGGCCCA	TGATGTCAGA	GCTGGAAGAG	AGGGCACGTC	3540
55	AGCAGAGGGG	CCACCTCCAT	TTGCTGGAGA	CAAGCATAGA	TGGGATTCTG	GCTGATGTGA	3600
	AGAACTTGGA	GAACATTAGG	GACAACCTGC	CCCCAGGCTG	CTACAATACC	CAGGCTCTTG	3660
	AGCAACAGTG	AAGCTGCCAT	AAATATTTCT	CAACTGAGGT	TCTTGGGATA	CAGATCTCAG	3720
	GGCTCGGGAG	CCATGTCATG	TGAGTGGGTG	GGATGGGGAC	ATTTGAACAT	GTTTAATGGG	3780
	TATGCTCAGG	TCAACTGACC	TGACCCCAT	CCTGATCCCA	TGGCCAGGTG	GTTGTCTTAT	3840
60	TGCACCATAC	TCCTTGCTTC	CTGATGCTGG	GCAATGAGGC	AGATAGCACT	GGGTGTGAGA	3900
	ATGATCAAGG	ATCTGGACCC	CAAAGAATAG	ACTGGATGGA	AAGACAAACT	GCACAGGCAG	3960
	ATGTTTGCCCT	CATAATAGTC	GTAAGTGGAG	TCCTGGAATT	TGGACAAGTG	CTGTTGGGAT	4020
	ATAGTCAACT	TATTCTTTGA	GTAATGTGAC	TAAAGGAAAA	AACTTTGACT	TTGCCCAGGC	4080
	ATGAAATTCT	TCCTAATGTC	AGAACAGAGT	GCAACCCAGT	CACACTGTGG	CCAGTAAAAT	4140
65	ACTATTGCCT	CATATTGTCC	TCTGCAAGCT	TCTTGCTGAT	CAGAGTTCCCT	CCTACTTACA	4200
	ACCCAGGGTG	TGAACATGTT	CTCCATTTTC	AAGCTGGAAG	AAGTGAGCAG	TGTTGGAGTG	4260
	AGGACCTGTA	AGGCAGGCCC	ATTGAGGCTT	ATGGTGCTTG	CTGGTGCCTG	CCACCTTCAA	4320
	GTCTCTGGAC	TGGGCATGAC	ATCCTTTCTT	TAAATGATGC	CATGGCAACT	TAGAGATTGC	4380
	ATTTTTTATTA	AAGCATTTC	TACCAGCAAA	GCAATGTTG	GGAAAGTATT	TACTTTTTTCG	4440
70	GTTTTCAAAGT	GATAGAAAAG	TGTGGCTTGG	GCATTGAAAG	AGGTAAAATT	CTCTAGATTT	4500
	ATTAGTCCTA	ATTCAATCCT	ACTTTTCGAA	CACCAAAAAT	GATGCGCATC	AATGTATTTT	4560
	ATCTTATTTT	CTCAATCTCC	TCTCTCTTTC	CTCCACCCAT	AATAAGAGAA	TGTTCTTACT	4620
	CACACTTCAG	CTGGGTCCAC	TCCATCCCTC	CATTATCCCT	TCCATCCATC	TTTCCATCCA	4680
	TTACCTCCAT	CCATCCTTCC	AACATATATT	TATTGAGTAC	CTACTGTGTG	CCAGGGGCTG	4740
75	GTGGGACAGT	GGTGACATAG	TCTCTGCCCT	CATAGAGTTG	ATTGTCTAGT	GAGGAAGACA	4800
	AGCATTTTTTA	AAAAATAAAT	TTAAACTTAC	AAACTTTGTT	TGTCACAAGT	GGTGTTTATT	4860
	GCAATAACCG	CTTGTTTTCG	AACCTCTTTG	CTCAACAGAA	CATATGTTGC	AAGACCCTCC	4920
	CATGGGGGCA	CTTGAGTTT	GGCAAGGCTG	ACAGAGCTCT	GGGTTGTGCA	CATTTCTTTG	4980
	TATCCAGAGT	GCTACTCTGT	GCCTTTCTAC	AACTGATTGC	AACAGACTGT	TGAGTTATGA	5040
80	CAACACCAGT	GGGAATTGCT	GGAGGAACCA	GAGGCACTTC	CACCTTGGCT	GGGAAGACTA	5100
	TGGTGCTGCC	TTGCTTCTGT	ATTTCTCTGG	ATTTCTCTGA	AAGTGTTTT	AAATAAAGAA	5160
	CAATTGTTAG	ATGCC					

Seq ID NO: 220 Protein sequence:  
Protein Accession #:NP\_005553

1	11	21	31	41	51



	MPALWLGCCD	CFSLLLLPAAR	ATSRREVCDC	NGKSRQCIFD	RELHRQTGNG	FRCLNCNDNT	60
	DGIHCEKCKN	GFYRHRERDR	CLPCNCNSKG	SLSARCDNSG	RCCKKPGVTG	ARCDRCCLPGF	120
	HMLTDAGCTQ	DQRLLDKCD	CDPAGIAGPC	DAGRCVCKPA	VTGERCDRCR	SGYVNLDDGN	180
5	PEGCTQCFCY	GHSASCRSSA	EYSVHKITST	FHQDVGWKA	VQRNGSPAKL	QWSQRHQDVF	240
	SSAQRLDPVY	FVAPAKFLGN	QQVSYGQSL	FDYRVDGRGR	HPSAHDVILE	GAGLRITAPL	300
	MPLGKTLPCG	LTKTYTFRLN	EHPSNNWSPQ	LSYFEYRRL	RNLTLALRIR	TYGEYSTGYI	360
	DNVTLLISARP	VSGAPAPWVE	QCICPVGYKG	QFCQDCASGY	KRDSARLGPF	GTICPCNCQG	420
	GGACDPDTGD	CYSGDENPDI	ECADCPIGFY	NDPHDPRSC	PCPCHNGFSC	SVMPEEEVV	480
10	CNNCPFGVTG	ARCELCADGY	FGDPFGEHGP	VRPCQPCQCN	NNVDPSASGN	CDRLTGRCCLK	540
	CIHNTAGIYC	DQCKAGYFGD	PLAPNPADKC	RACNCPMGS	EPVGCSDGT	CVCKPGFGGP	600
	NCEHGAFCSP	ACYNQVKIQM	DQFMQQLQRM	EALISKAQGG	DGVVPDTELE	GRMQQAQAL	660
	QDILRDAQIS	EGASRLGLQ	LAKVRSQENS	YQSRLLDLKM	TVERVRALGS	QYQNRVRDTH	720
	RLITQMLSL	AESEASLGNT	NIPASDHVVG	PNGFKSLAQE	ATRLAESHVE	SASNMEQLTR	780
15	ETEDYSKQAL	SLVRKALHEG	VSGSGSPDG	AVVQGLVEKL	EKTKSLAQQL	TREATQAEIE	840
	ADRSYQHSRL	LLDSVSRLOQ	VSDQSFQVEE	AKRIKQKADS	LSTLVTRHMD	EFKRTQKNLG	900
	NWKEEAQQLL	QNGKSGREKS	DQLLSRANLA	KSRAQEALSM	GNATFYEVE	ILKNLREFDL	960
	QVDNRKAEAE	EAMKRLSYIS	QKVSADSDKT	QQAERALGSA	AADAQRAKNG	AGEALEISSE	1020
	IEQEIGSLNL	EANVTADGAL	AMEKGLASLK	SEMREVEGEL	ERKELEFDTN	MDAVQMVITE	1080
20	AQKVDTRAKN	AGVTIQDTLN	TLDGLHLMD	QPLSVDEEGL	VLLEQKLSRA	KTQINSQLRP	1140
	MMSELEERAR	QQRGHLHLE	TSIDGILADV	KNLENIRDNL	PPGCYNTQAL	EQQ	

Seq ID NO: 221 DNA sequence  
Nucleic Acid Accession #: NM\_016529  
Coding sequence: 13-1854

25	1	11	21	31	41	51	
	GTCAAGAAAA	GAATGTCTGT	AATTGTTCGA	ACTCCTTCAG	GACGACTTCG	GCTTTACTGT	60
30	AAAGGGGCTG	ATAATGTGAT	TTTTTGAGAGA	CTTTCAAAAG	ACTCAAAATA	TATGGAGGAA	120
	ACATTATGCC	ATCTGGAATA	CTTTGCCACG	GAAGGCTTGC	GGACTCTCTG	TGTGGCTTAT	180
	GCTGATCTCT	CTGAGAATGA	GTATGAGGAG	TGGCTGAAAG	TCTATCAGGA	AGCCAGCACC	240
	ATATTGAAGG	ACAGAGCTCA	ACGGTTGGAA	GAGTGTTACG	AGATCATTGA	GAAGAATTTG	300
	CTGCTACTTG	GAGCCACAGC	CATAGAAGAT	CGCCTTCAAG	CAGGAGTTCC	AGAAACCATC	360
35	GCAACACTGT	TGAAGGCAGA	AATTAAAATA	TGGGTGTTGA	CAGGAGACAA	ACAAGAAACT	420
	GCGATTAATA	TAGGGTATTC	CTGCCGATTG	GTATCGCAGA	ATATGGCCCT	TATCCTATTG	480
	AAGGAGGACT	CTTTGGATGC	CACAAGGGCA	GCCATTACTC	AGCACTGCAC	TGACCTTGGG	540
	AATTTGCTGG	GCAAGGAAAA	TGACGTGGCC	CTCATCATCG	ATGGCCACAC	CCTGAAGTAC	600
	GCGCTCTCCT	TCGAAGTCCG	GAGGAGTTTG	CTGGATTGG	CACCTCTCGT	CAAAGCGGTC	660
40	ATATGCTGCA	GAGTGCTCTC	TCTGCAGAAC	TCTGAGATAG	TGGATGTGGT	GAAGAAGCGG	720
	GTGAAGGCCA	TCACCCTCGC	CATCGGAGAC	GGCGCCAACG	ATGTCGGGAT	GATCCAGACA	780
	GCCACGTTGG	GTGTGGGAAT	CAGTGGGAAT	GAAGGCATGC	AGGCCACCAA	CAACTCGGAT	840
	TACGCCATCG	CACAGTTTTT	CTACTTAGAG	AAGCTTCTGT	TGGTTCATGG	AGCCTGGAGC	900
	TACAACCGGG	TGACCAAGTG	CATCTTGTAC	TGCTTCTATA	AGAACGTGGT	CCTGTATATT	960
45	ATTGAGCTTT	GGTTCGCCTT	TGTTAATGGA	TTTCTGGGCG	AGATTTTATT	TGAACGTTGG	1020
	TGCATCGGCC	TGTACAATGT	TGTTTTCACC	GCTTTGCCGC	CCTTCACCTC	GGGAATCTTT	1080
	GAGAGGTCTT	GCACTCAGGA	GAGCATGCTC	AGGTTTCCCC	AGCTCTACAA	AATCACCCAG	1140
	AATGGCGAAG	GCTTCAACAC	AAAGGTTTTT	TGGGGTCACT	GCATCAACGC	CTTGGTCCAC	1200
	TCCCTCATCC	TCTTCTGGTT	TCCCATGAAA	GCTCTGGAGC	ATGATACTGT	GTGTGACAGT	1260
50	GGTCATGCTA	CCGACTATTT	ATTTGTTGGA	AATATTGTTT	ACACATATGT	TGTTGTTACT	1320
	GTTTGCTCTG	AAGCTGGTTT	GGAGACCACA	GCTTGGACTA	AATTCAGTCA	TCTGGCTGTC	1380
	TGGGGAAGCA	TGCTGACCTG	GCTGGTGTCT	TTTGCCATCT	ACTCGACCAT	CTGGCCCAAC	1440
	ATTCCCATTG	CTCCAGATAT	GAGAGGACAG	GCAACTATGG	TCCTGAGCTC	CGCACACTTC	1500
	TGGTTGGGAT	TATTTCTGGT	TCCTACTGCC	TGTTTGATTG	AAGATGTGGC	ATGGAGAGCA	1560
55	GCCAAGCACA	CCTGCAAAAA	GACATTGCTG	GAGGAGGTGC	AGGAGCTGGA	AACCAAGTCT	1620
	CGAGTCCTGG	GAAAAGCGGT	GCTGCGGGAT	AGCAATGGAA	AGAGGCTGAA	CGAGCGCGAC	1680
	CGCCTGATCA	AGAGGCTGGG	CCGGAAGACG	CCCCGACGCG	TGTTCCGGGG	CAGCTCCCTG	1740
	CAGCAGGGCG	TCCCGCATGG	GTATGCTTTT	TCTCAAGAAG	AACACGGAGC	TGTTAGTCAG	1800
	GAAGAAGTCA	TCCGTGCTTA	TGACACCACC	AAAAAGAAAT	CCAGGAAGAA	ATAAGACATG	1860
60	AATTTTCCTG	ACTGATCTTA	GGAAAGAGAT	TCAGTTTGTT	GCACCCAGTG	TTAACACATC	1920
	TTTGTCAGAG	AAGACTGGCG	TCCAAGGCCA	AAACACCAGG	AAACACATTT	CTGTGGCCTT	1980
	AGTTAAGCAG	TTTGTTAGTT	ACATATTTCC	TCGCAAACTT	GGAGTGCAGA	CCACAGGGGA	2040
	AGCTATCTTT	GCCCTCCCAA	CTCGTCTGCA	GTGCTTAGCC	TAACTTTTGT	TTATGTCTGT	2100
65	ATGAAGCATT	CAACTGTGCT	CTGTGAGGTC	TCAATTAATA	AACATTATGT	TTACCAATA	2160
	AGAAAAA	AAAAA					

Seq ID NO: 222 Protein sequence:  
Protein Accession #: NP\_057613

70	1	11	21	31	41	51	
	MSVIVRTPSG	RLRLYCKGAD	NVIFERLSKD	SKYMEETLCH	LEYFATEGLR	TLCVAYADLS	60
	ENEYEELKV	YQEAETILKD	RAQRLEECYE	IEKNLLLLG	ATAIEDRLQA	GVPETIATLL	120
	KABIKIWLVT	GDKQETAINI	GYSCRLVSQN	MALILLKEDS	LDATRAAITQ	HCTDLGNLLG	180
75	KENDVALIID	GHTLKYALSF	EVRRSFLDLA	LSCKAVICCR	VSPLQKSEIV	DVVKKRVKAI	240
	TLAIGDGAND	VGMQTAHV	VGISGNEGMQ	ATNNSDYAIA	QFSYLEKLL	VHGAWSYNRV	300
	TKCILYCFYK	NVVLYIIELV	FAFVNGFSGQ	ILNFERWCIGL	YNVIFTALPP	FTLGIFERS	360
	TQESMLRFPQ	LYKITQNGEG	FNTKVFNGHC	INALVHSLIL	FWFPMKALEH	DTVFDSEHAT	420
	DYLFVGNIVY	TYVVVTVCLK	AGLETTAWTK	FSHLAVWGSM	LTWLVEFGIY	STIWPTIPIA	480
80	PDMRGQATMV	LSSAHFWLGL	FLVPTACLE	DVAWRAAKHT	CKKTLLEEVQ	ELETKSRVLG	540
	KAVLRDSNGK	RLNERDRLIK	RLGRKTPPTL	FRGSSLLQGV	PHGYAFSQEE	HGAVSQBEVI	600
	RAYDTTKKKS	RKK					

Seq ID NO: 223 DNA sequence  
Nucleic Acid Accession #: BC017001  
Coding sequence: 1-394

85	1	11	21	31	41	51
----	---	----	----	----	----	----



5  
10  
15  
20  
25  
30  
35  
40  
45

AACGCTGGGC	AGGGCCGGCG	CGGGTCGGGG	GGCGCCCGAG	GGGCCC GGCG	CGAGCGGCGG	60
CGCGCAGGGC	GGCAGCATCC	ACTCGGGCCG	CATCGCCGCG	GTGCACAACG	TGCCGCTGAG	120
CGTGCTCATC	CGGCCGCTGC	CGTCCGTGTT	GGACCCCGCC	AAGGTGCAGA	GCCTCGTGGA	180
CACGATCCGG	GAGGACCCAG	ACAGCGTGCC	CCCCATCGAT	GTCTCTTGGA	TCAAAGGGGC	240
CCAGGGAGGT	GACTACTTCT	ACTCCTTTGG	GGGCTGCCAC	CGCTACGCGG	CCTACCAGCA	300
ACTGCAGCGA	GAGACCATCC	CCGCCAAGCT	TGTCCAGTCC	ACTCTCTCAG	ACCTAAGGGT	360
GTACCTGGGA	GCATCCACAC	CAGACTTGCA	GTAGCAGCCT	CCTTGGCACC	TGCTGCCACC	420
TTCAAGAGCC	CAGAAGACAC	ACCTGGCCTC	CAGCAGGCTG	GGCCATGCAG	AAGGGATAGC	480
AGGGGTGCAT	TCTCTTTGCA	CCTGGCGAGA	GGTCTGACT	CTGGGCACCC	CTCTCACCCG	540
CTACAAGGCC	TTGGACTCAC	TGTACAGTGT	GGGAGCCCCA	GTTCCACCT	CTGTGACAA	600
AGGATCATGG	CCTTACCCTT	GAAGCATTAC	CGAGAAGGAG	AACAGAGATG	GGCTTGAAGA	660
GCCACGTGCT	GCCGGCTCCA	AATTCCCAAG	GACAAGGATC	CCTCTGCATT	TTTGTCTATG	720
TAACTCTTA	TATGGACTAC	ATTGAGCTGC	AAGGAAAGGA	AAACCTTGAT	TGCAGTGGTT	780
TAAACAAACA	GAAGATTGTT	TTTCCACATA	GCATGGATTG	TGGAGATGGG	TGGCTAATGG	840
TATTGGTTCA	ACAATCCAC	GGAGGTAGGG	GTACCGTCTT	GGATCCTTTT	GCCTTAATCT	900
CAGTGCTCGT	TACTTCATGG	TCCCAAGATG	GCTGCTGTAT	CCCCAAGAAT	CATGTCTGCG	960
TTCAAGGAAG	GAGGGGTGGA	GGAAGAGGAA	GGGCCAAACT	AGCTGGACCC	GTCACCTTCT	1020
ATCAGAAAGT	AAAACCTCGT	CAGAAGTCTG	TTTCTGCTC	TCTCCCTCTG	CATATCTTCA	1080
CTTAGATGCC	CTTGGCCCGA	GCCAGCTACC	ATTGCACCTC	TAGCTGCAAA	CAAAGCTAAG	1140
ACAGCAGGGA	ACAGAATTGT	CATGGCTGAA	TAGACCAATC	GTGTTCCATC	TACTGAGACT	1200
GGCACACTGC	CTCCTGCAAT	AAAATGGGA	TCCCATTACC	AAGAGAGAAA	TGCAGAATTG	1260
TGTACCACTG	AGCTTTTGCT	GTGTAACAAA	CCATCCCCAA	ACTTGGCAGC	TAGAAACAAA	1320
CCCTGTATTT	TCCCAACAATC	CTATGGGTTG	GCAATTTGGG	CTGGGCTCAA	CAGGGCAGTT	1380
CTGCTGCTCA	CACCTGGGAT	CCCTCATGGA	GCTAAGGTCA	GCTGTTACCT	CAGCTGGGCC	1440
TGGATGCTCT	AGGATAGCCT	TACTCACTTG	CCTGGCAGGT	GACAGGCTGT	TGGCTGGAAT	1500
TGCTTGTTTC	TCTCCATGT	GGCCTCTCCA	GCAGGCTAGC	TCAGGCTTAT	TCACATGATG	1560
GCTTCAGGAT	TCCAAAGAGA	GTGAGAGTAG	AAGCTGAAAG	ACTTCTTGAG	TTCTTGCCCT	1620
GGAACCTGGG	CTAGGACAGT	GTCACCTCTG	CTAAGTTCTT	TTGGTCAGAG	CAAATCACAA	1680
GGCTTTACCC	AGATTCAAGG	GATGAGAAAC	AGACTACATG	TCTTGATGAG	GGGAACCACA	1740
AAGAGCTTGT	GGCCATTTT	CACCTATCAC	AAATAATTTT	GGATGGGTAT	TTATTTGGAT	1800
AAAGGTATTT	CCCTCTTCCC	CCTTTCTCTC	TGTCTCATGG	GGCCTCACTC	TGCCAAGTTG	1860
GAAGGCACCTA	AGACATTGTC	CTGGCCCTCA	GGGTCTAGGG	GAAGAGGTGT	TGGGGCAGGA	1920
AGTGAGTCTC	TCCATGGGCT	GGACCCACTG	TAGTAGGAGT	GCCTCCTTGT	CTGCACTGCT	1980
GGTATGGGGT	TAGGCCAGGT	AGGACATTCC	AGAGGGGCTT	CTGAAAACCA	AGAGTCCCTG	2040
GGGAAAGGGA	ACAGAGTAAG	GCAGGCCTTG	TTCTCACTGC	CCTCTAAGGG	AACTTGGTCA	2100
CTCGGCACCTT	TTAAGCCTCA	GTTTCTCCAG	TTCAATAATA	AGGACAAGAG	CTTTTCCCAT	2160
GCATTCTCTT	TCCCCGGGAA	AGTTGACTGA	GGTGACCACT	AATAGAATTG	AAAAGGGAGA	2220
GTGTCTTCAG	TGCAATGTGG	CATCCTGGAT	TGGGTCTTGG	AACAAAAACA	GGACATTAGT	2280
GGGAAATTG	GAAATCTGAA	AAAAGTCTGA	ATTTTAGTTA	ATATACCAAT	TTCAGTCTCT	2340
TGGTTTTGAC	AGATGTACCA	TGGTGATGTA	AGATGTTGAC	CTTGGGGTAG	GCTGGGTGAA	2400
GGGTATACAG	GAACTCTTTG	TACTATCTCT	GCAACTTCTC	TGTAAATCTA	GTATCATTCC	2460
AAAATAAAG	TTTATTTAAT	TTAAAAAAA	AAAAAAA	AA		

Seq ID NO: 224 Protein sequence:  
Protein Accession #: AAH17001.1

50

1	11	21	31	41	51	
TLGRAGAGRG	APEGPGPSGG	AQGGSIHSGR	IAAVHNVPLS	VLIRPLPSVL	DPAKVQSLVD	60
TIREDPDSVP	PIDVLWIKGA	QGGDYFYFSFG	GCHRYAAYQQ	LQRETIPAKL	VQSTLSDLRV	120
YLGASTPDLQ						

55  
Seq ID NO: 225 DNA sequence  
Nucleic Acid Accession #: NM\_021048  
Coding sequence: 1..1110

60  
65  
70  
75  
80

1	11	21	31	41	51	
ATGCCTCGAG	CTCCAAAGCG	TCAGCGCTGC	ATGCCTGAAG	AAGATCTTCA	ATCCCAAAGT	60
GAGACACAGG	GCCTCGAGGG	TGCACAGGCT	CCCCTGGCTG	TGGAGGAGGA	TGCTTCATCA	120
TCCACTTCCA	CCAGCTCCTC	TTTTCCATCC	TCTTTTCCCT	CCTCCTCCTC	TTCTCCTCC	180
TCCTCCTGCT	ATCCTCTAAT	ACCAAGCACC	CCAGAGGAGG	TTTCTGCTGA	TGATGAGACA	240
CCAAATCCTC	CCCAGAGTGC	TCAGATAGCC	TGCTCCTCCC	CCTCGGTCGT	TGCTTCCCTT	300
CCATTAGATC	AATCTGATGA	GGGCTCCAGC	AGCCAAAAGG	AGGAGAGTCC	AAGCACCCTA	360
CAGGTCTGCT	CAGACAGTGA	GTCTTTACCC	AGAAGTGAGA	TAGATGAAAA	GGTACTGAT	420
TTGGTGAGT	TTCTGCTCTT	CAAGTATCAA	ATGAAGGAGC	CGATCACAAA	GGCAGAAATA	480
CTGGAGAGTG	TCATAAAAAA	TTATGAAGAC	CACTTCCCTT	TGTTGTTTAG	TGAAGCCTCC	540
GAGTGATGCT	TGCTGGTCTT	TGGCATTGAT	GTAAGGAAG	TGGATCCCAC	TGGCCACTCC	600
TTTGTCCTTG	TCACCTCCCT	GGGCCTCACC	TATGATGGGA	TGCTGAGTGA	TGTCCAGAGC	660
ATGCCCAAGA	CTGGCATTCT	CATACCTTATC	CTAAGCATAA	TCTTCATAGA	GGGCTACTGC	720
ACCCCTGAGG	AGGTCATCTG	GGAAGCAGTG	AATATGATGG	GGCTGTATGA	TGGGATGGAG	780
CACCTCATTT	ATGGGGAGCC	CAGGAAGCTG	CTACCCCAAG	ATTGGGTGCA	GGAAACTAC	840
CTGGAGTACC	GGCAGGTGCC	TGGCAGTGAT	CCTGCACGGT	ATGAGTTTCT	GTGGGGTCCA	900
AGGGCTCATG	CTGAAATTAG	GAAGATGAGT	CTCCTGAAAT	TTTTGGCCAA	GGTAAATGGG	960
AGTGATCCAA	GATCCTTCCC	ACTGTGGTAT	GAGGAGGCTT	TGAAAGATGA	GGAAGAGAGA	1020
GCCAGGACA	GAATTGCCAC	CACGATGAT	ACTACTGCCA	TGGCCAGTGC	AAGTTCTAGC	1080
GCTACAGGTA	GCTTCTCCTA	CCCTGAATAA				

Seq ID NO: 226 Protein sequence:  
Protein Accession #: NP\_066386

85

1	11	21	31	41	51	
MPRAPKRQRC	MPEEDLQSQS	ETQGLEGAQA	PLAVEEDASS	STSTSSSFPS	SFPSSSSSSS	60



SSCYPLIPST PEEVSADDET PNPPQSAQIA CSSPSVVASL PLDQSDEGSS SQKEESPSTL 120  
 QVLPDSESLP RSEIDEKVTD LVQFLLFKYQ MKEPITKAEI LESVIKNYED HFPLLFSEAS 180  
 ECMLLVFGID VKEVDPTGHS FVLVTSLGLT YDGM LSDVQS MPKTGILILI LSIIFIEGYC 240  
 TPPEVIWEAL NMMGLYDGME HLIYGEPRKL LTQDWVQENY LEYRQVPGSD PARYEFLWGP 300  
 RAHAEIRKMS LLKFLAKVNG SDPRSFLWY BEALKDEEER AQDRIATTD TTMASASSS 360  
 ATGSFSYPE

Seq ID NO: 227 DNA sequence  
 Nucleic Acid Accession #: NM\_005025.1  
 Coding sequence: 82-1314

1 11 21 31 41 51  
 | | | | |  
 GCGGAGCACA GTCCGCCGAG CACAAGCTCC AGCATCCCGT CAGGGGTTGC AGGTGTGTGG 60  
 GAGGCTTGAA ACTGTTACAA TATGGCTTTC CTTGGACTCT TCTCTTTGCT GGTTCCTGCA 120  
 AGTATGGCTA CAGGGGCCAC TTTCCCTGAG GAAGCCATTG CTGACTTGTC AGTGAATATG 180  
 TATAATCGTC TTAGAGCCAC TGGTGAAGAT GAAAATATTC TCTTCTCTCC ATTGAGTATT 240  
 GCTCTTGCAA TGGGAATGAT GGAACCTGGG GCCCAAGGAT CTACCCAGAA AGAAATCCGC 300  
 CACTCAATGG GATATGACAG CCTAAAAAAT GGTGAAGAAT TTTCTTTCTT GAAGGAGTTT 360  
 TCAAACATGG TAACTGCTAA AGAGAGCCAA TATGTGATGA AAATTGCCAA TTCCTTGTTT 420  
 GTGCAAAATG GATTTCATGT CAATGAGGAG TTTTTCGAAA TGATGAAAAA ATATTTTAAT 480  
 GCAGCAGTAA ATCATGTGGA CTTCACTCAA AATGTAGCCG TGGCCAACTA CATCAATAAG 540  
 TGGGTGGAGA ATAACACAAA CAATCTGGTG AAAGATTGTT TATCCCCAAG GGATTTTGAT 600  
 GCTGCCACTT ATCTGGCCCT CATTAATGCT GTCTATTTCA AGGGGAACTG GAAGTCGCAG 660  
 TTTAGGCCTG AAAATACTAG AACCTTTTCT TTCATAAAG ATGATGAAAG TGAAGTCCAA 720  
 ATTCCAATGA TGTATCAGCA AGGAGAATTT TATTATGGGG AATTTAGTGA TGGCTCCAAT 780  
 GAAGCTGGTG GTATCTACCA AGTCCTAGAA ATACCATATG AAGGAGATGA AATAAGCATG 840  
 ATGCTGGTGC TGTCCAGACA GGAAGTTTCT CTTGCTACTC TGGAGCCATT AGTCAAAGCA 900  
 CAGCTGGTTG AAGAATGGGC AAACCTCTGTG AAGAAGCAAA AAGTAGAAGT ATACCTGCCC 960  
 AGGTTACAGT TGAACAGGA AATTGATTTA AAAGATGTTT TGAAGGCTCT TGAATAAAT 1020  
 GAAATTTTCA TCAAAGATGC AAATTTGACA GGCTCTCTCTG ATAATAAGGA GATTTTCTT 1080  
 TCCAAAGCAA TTCACAAGTC CTTCTAGAGT GTTAATGAAG AAGGCTCAGA AGCTGCTGCT 1140  
 GTCTCAGGAA TGATTGCAAT TAGTAGGATG CTTGTGCTGT ATCCTCAAGT TATTGTCGAC 1200  
 CATCCATTTT TCTTTCTTAT CAGAAAACAGG AGAACTGGTA CAATTCTATT CATGGGACGA 1260  
 GTCATGCATC CTGAACAAT GAACACAAGT GGACATGATT TCGAAGAACT TTAAGTTACT 1320  
 TTATTTGAAT AACAAGGAAA ACAGTAACTA AGCACATTAT GTTTGCAACT GGTATATATT 1380  
 TAGGATTTGT GTTTTACAGT ATATCTTAAAG ATAATATTTA AAATAGTTCC AGATAAAAAAC 1440  
 AATATATGTA AATTATAAGT AACTTGTCAA GGAATGTTAT CAGTATTAAG CTAATGGTCC 1500  
 TGTTATGTCA TTGTGTTTGT GTGCTGTTGT TTAAATAAAA AGTACCTATT GAACATGTG

Seq ID NO: 228 Protein sequence:  
 Protein Accession #: NP\_005016.1

1 11 21 31 41 51  
 | | | | |  
 MAFLGLFSLV VLQSMATGAT FPPEAIADLS VNMYNRLRAT GEDENILFSP LSIALAMGMM 60  
 ELGAQGSTQK EIRHSMGYDS LKNGEESFSL KEFSNMVTAK ESQYVMKIAN SLFVQNGFHV 120  
 NEEFLQMMKK YFNAAVNHDV FSQNVAVANY INKWVENNTN NLVKDLVSPR DFDAATYLAL 180  
 INAVYFKGNW KSQFRPENTR TFSFTKDDDES EVQIPPMYQQ GEFYGEFSD GSNEAGGIYQ 240  
 VLEIPYEGDE ISMMLVLSRQ EVPLATLEPL VKAQLVEEWA NSVKKQKVEV YLPRFTVEQE 300  
 IDLKDVLKAL GITEIFIKDA NLTGLSDNKE IFLSKAIHKS FLEVNEEGSE AAASVGMIAI 360  
 SRMAVLYPQV IVDHPFFFLI RNRRTGTILF MGRVMHPETM NTSGHDFEEL

Seq ID NO: 229 DNA sequence  
 Nucleic Acid Accession #: NM\_003695  
 Coding sequence: 12-398

1 11 21 31 41 51  
 | | | | |  
 CGACATCAGA GATGAGGACA GCATTGCTGC TCCTTGACAGC CCTGGCTGTG GCTACAGGGC 60  
 CAGCCCTTAC CCTGCGCTGC CACGTGTGCA CCAGCTCCAG CAACTGCAAG CATTCTGTGG 120  
 TCTGCCCCGGC CAGCTCTCGC TTCTGCAAGA CCACGAACAC AGTGGAGCCT CTGAGGGGGA 180  
 ATCTGGTGAA GAAGGACTGT GCGGAGTCGT GCACACCCAG CTACACCCTG CAAGGCCAGG 240  
 TCAGCAGCGG CACCAGCTCC ACCCAGTCTG GCCAGGAGGA CCTGTGCAAT GAGAAGCTGC 300  
 ACAACGCTGC ACCCACCCTG ACCGCCCTCG CCCACAGTGC CCTCAGCCTG GGGCTGGCCC 360  
 TGAGCCTCCT GGCCGTCATC TTAGCCCCCA GCCTGTGACC TTCCCCCAG GGAAGGCCCC 420  
 TCATGCCTTT CCTTCCCTTT CTCTGGGGAT TCCACACCTC TCTTCCCCAG CCGGCAACGG 480  
 GGGTGCCAGG AGCCCCAGGC TGAGGGCTTC CCCGAAAGTC TGGGACCAGG TCCAGGTGGG 540  
 CATGGAATGC TGATGACTTG GAGCAGGCCC CACAGACCCC ACAGAGGATG AAGCCACCCC 600  
 ACAGAGGATG CAGCCCCCAG CTGCATGGAA GGTGGAGGAC AGAAGCCCTG TGGATCCCCG 660  
 GATTTACAC TCTTCTGTT TTGTTGCCGT TTATTTTGTA CTCAAATCTC TACATGGAGA 720  
 TAAATGATTT AAAC

Seq ID NO: 230 Protein sequence:  
 Protein Accession #: NP\_003686

1 11 21 31 41 51  
 | | | | |  
 MRTALLLLAA LAVATGPALT LRCHVCTSSS NCKHSVVCPA SSRFCKTTNT VEPLRGNLVK 60  
 KDCAESCTPS YTLQGVSSG TSSTQCCQED LCNEKLHNAA PTRTALAHS LSLGLALSL 120  
 AVILAPSL

Seq ID NO: 231 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 126-752



1	11	21	31	41	51	
CCGGGCAGGT	GGCTCATGCT	CGGGAGCGTG	GTTGAGCGGC	TGGCGCGGTT	GTCTTGGAGC	60
AGGGGCGCAG	GAATTCTGAT	GTGAACTAA	CAGTCTGTGA	GCCCTGGAAC	CTCCACTCAG	120
AGAAGATGAA	GGATATCGAC	ATAGGAAAAG	AGTATATCAT	CCCCAGTCCT	GGGTATAGAA	180
GTGTGAGGGA	GAGAACCAGC	ACTTCTGGGA	CGCACAGAGA	CCGTGAAGAT	TCCAAGTTCA	240
GGAGAACTCG	ACCGTTGGAA	TGCCAAGATG	CCTTGGAAAC	AGCAGCCCGA	GCCGAGGGCC	300
TCTCTCTTGA	TGCCTCCATG	CATTCTCAGC	TCAGAATCCT	GGATGAGGAG	CATCCCAAGG	360
GAAAGTACCA	TCATGGCTTG	AGTGCTCTGA	AGCCCATCCG	GACTACTTCC	AAACACCAGC	420
ACCCAGTGGA	CAATGCTGGG	CTTTTTCCT	GTATGACTTT	TTCGTGGCTT	TCTTCTCTGG	480
CCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
AAGTTGGGCC	AGACGCTGCT	TCCCTGCGAA	GGGTGTGTG	GATCTTCTGC	CGCACCAGGC	660
TCATCCTGTC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAAATT	720
TTCAGGATGG	CTGTATTCTG	CGGTCAGAAT	GAGAGAGTCA	AGCTGGGCAG	AATCTCTCGC	780
CAAGAGTTCA	GCCTTCCTTT	GGAGACTGCT	CCATCAGTGC	CGAGGTGTGT	GGGAACAGGC	840
TTCAGTGCAC	CGCCATCTTA	CTGAGTTGCT	TCACGTGAGG	AAAAGGGGGC	TTTGGCCCTG	900
TGACTCAGTT	CCACATTTTG	GATTGCATAC	TGGAAGAA	GCCAATCTTC	TTGCTAGTAA	960
ACCGCAACC	CGGCTGTATA	CAGTGGTGAC	CCAAGCAATG	GATATAAACC	TAAAAATCTG	1020
AGGGAGGGGA	GAGGTGGAAT	ACAGTAGTTC	TTGGAATCTG	AAGTCTCCTA	TTTGATCAGG	1080
TTATTTCTTG	GGACTTGGCA	AAAACTGAT	TGGTGGGGAT	CTCCTAGGAC	CTAGTGGACA	1140
TCTGGTATTA	ATTTAATCTC	AGGAAAAACA	AGAAATTAAC	CCAGAGAGAG	TCTGGGTTTT	1200
GGAATTCAGC	GTAGCTACCT	CCAGACCGTG	GTGTCTGGCC	TCCATTTTGT	TCTGTCAATC	1260
AGCTCTGACT	TACAGCTGCA	GTCACCTTTG	CTATAAGGCA	CCTGGGTAGA	AGGGTGGATG	1320
GGCTTCACAT	CAATTTTTTT	CTTCCTTAG	GGTGGGGGAT	TGGTTTGGCT	TTCTTTTGTT	1380
GTGGTTTTTT	GTTTTATTTT	TGTCAGATT	GATTTTTTGA	TGCAAGGACT	TGAAAAGACC	1440
CAGAAGGATG	CCACCAGTTT	TTCCTTGAGG	CCTAGGATTT	TTTATTTCTGT	CCCGAGCAGA	1500
GGTAATTCCT	CACAACTTAG	TGCACCAGTA	GCACCAGCCA	TTTTGAGCAG	AGTACCTCTT	1560
TGGGGAGCTT	TTCGTTTTGT	TTTGTTTTGA	ATTCTCTTTC	CTTAGCAGCA	AGGTCTTTTT	1620
TCCTAGAGAA	TCTACTCCGT	TGCAGAATCA	TTGCAACCTC	AGGAGCCCTC	ACTGATTGAG	1680
TGCTGTCAGC	CTGATATACT	ACTTTGGACT	CTGGAAACAG	ATATGGGTTT	TATTTCTCTAT	1740
TTCTACTGTG	TGTCGTTAAA	CAACCGTCGG	AGACCAAGATG	ACCTGTTAGA	TGGCTAGTCC	1800
TGTATAACTC	GACTCTGTAT	GTTTCAATGT	ATGTTACTGC	AATGCTTCAC	CTGCTGTACA	1860
GTGTTTGTGA	GATGCTCTTT	GAAGATGGTA	CTTTTATATT	T		

Seq ID NO: 232 Protein sequence:  
Protein Accession #: Eos sequence

1	11	21	31	41	51	
MKDIDIGKEY	IIPSPGYRSV	RERTSTSGTH	RDREDSKFRF	TRPLECQDAL	ETAARAEGLS	60
LDASMHSQLR	ILDEEHPKKG	YHHGLSALKP	IRTTSKHQHP	VDNAGLFSCM	TFSWLSSLAR	120
VAHKKGELSM	EDVWSLSKHE	SSDVNCRRL	RLWQEELNEV	GPDAASLRRV	VWIFCRTRLI	180
LSIVCLMITQ	LAGFSGPNFQ	DGCILRSE				

Seq ID NO: 233 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1	11	21	31	41	51	
TTTTAATGGT	GCTCATATAT	ACTGTATTTT	TTGTTGTTTA	GTTTTACTTA	TTGAGAGTGT	60
CACAACATGA	ATCACATAAT	CATGATTTTT	TTTTTTTACT	TTTACTCCCC	AAATTATTCA	120
TGTTTCTTAG	ATCGTAGTCA	TTGAGAAGTC	CCAATAACTC	TAAACTTTTG	AGTTATAACG	180
TAGTAAACTT	CTCTTTTCATC	TTTGTGTTAG	CTCTGTAGTC	TTAACCTGGA	TTTAAATTTT	240
TTTGTTTCCA	AAGTCACAAAT	TGAATATATC	TTAGATACCT	TAAGCCACTG	AATTCAGTTC	300
TGTTTGACTG	AAAGCAAAAC	AACGTGACAG	TTTATTTTCA	AACACTAACT	TCTTGATATT	360
TTGTTATGGT	ATATCTTTTT	ATTAAATATT	TATTTTGACT	AAGCTTTCAT	AAAATATTTG	420
AAGCTATTTT	AATCATCAAG	TATGAAAAAC	AAATTACTAT	TGCATTTTCC	TATATATGCA	480
TATATTATGG	ATTAACCAGA	ATTGTATCAT	TTTTGGCCTA	ATGTCTGGAT	ATAAAAGATA	540
ATTAGCCTAC	TATAGTATTA	ATAAATTTTT	CAGTTGGTTT	GGGCAAATTT	AAACCTGAAA	600
AATAGGTTAA	AAAGTAGTTA	CAAAATTAAAC	TTACTAATTT	ATACCTGATT	TTTTTTCTTG	660
AATTAAGTA	CATTTTAAAT	GAGCTTTATA	ATACCTTAAA	AAGTTGGTTC	TAATTTAAAA	720
TATGAAAGCT	CTGGCTATCA	TCCTGGGATA	GTAATTTCTA	ATTATATAGT	ATTCAAAC	780
TATATATTTT	TTAGTTCCCT	TGAGATAACT	AAATTTCTAAT	TATATATGTT	TCAAAAACCA	840
TATCCTGTAT	TTTTTTTAAAG	AATTGTTTTA	TAAATAGGTC	ATAAGATACA	AGGTCTGCAT	900
TAGAAGACCC	ACTCTTACTA	GGTTCCTTAA	GGATCTGCCA	TAGATTTTTT	TTTTTTTTTT	960
TTTTTTTTTAG	GTAGTTTAAA	GCAAGCACTG	ATACCAGTGG	GAGTTGGTCT	TGATCTAGGA	1020
GATTCTGTTA	AGCATCCAAA	AACAATGCCT	AATTTTCAGT	CTTAGGTTAT	GGCTTGAGAC	1080
TCCAGATAAA	AGATGGAGAA	TACCTCATGT	ACTGTGACTT	GAAAATGAAT	TCTTAAATTT	1140
CTTAGGCTCT	CTCCATGTAT	CTTCTTAAAG	GAAAAGTTTC	TGAGTGTGAT	CTCTCTTTTG	1200
CCATAGTATC	AAGTGGAGGG	TAGTTCAGAA	AAGTTAATAG	GAAATCTTTT	GTGACAGCAG	1260
ACTATAATAG	AAGTTTGAGT	AATATTTTAA	TAAATTTATA	TAATTCAAAT	GATAAAAAATG	1320
TATCAATGTT	ATCCAATGAT	TTTTATTAAA	AAATTACCTT	ATTATTAGAA	CTGTGCCTAT	1380
TACATAAAAA	GTGCTCATGT	ATTTGAATTT	TAAATAATTT	ATTTAAATCA	AGACCACCAT	1440
AAGTCATTAA	TAATTTAATA	ATTGTTTTAA	ATCAGTGGTT	TTCAACCCTC	ACTTCATATT	1500
AGAAATCATCT	GAGGACTTTT	AATATGGAAT	CCACCTCATA	ACAATTAAGT	CTAAATTTCT	1560
GGAAGATGGA	GCCATGCTTG	TTTTTCCAAA	AGCTCTTTGA	GTGATTCTAA	TTTGTAGTCA	1620
GAGTTGAAGA	CCACTGCTCT	AAATTAGTGC	AGGAAAATGC	TTTTATTCTT	CCCATGTTAA	1680
CTTTTAAAC	TAGTAATGTA	CCGATTAAAG	TTTTGATGGT	TTAAATTCCA	CTAAAGAACA	1740
TATTCTTCTA	ATAACTAGCA	TTTATTACAT	GAAATTTAAG	AGTTTAAGTT	CCATCAAACCT	1800
AGCCCTTGTTG	TAAGATTATT	ATTTCTTCTC	TATAACTTCA	AAATAGATAT	TTCAATCAAA	1860
CTGTTTCAGGT	GAGAAAACAT	AATGGATTTT	TTTTTTTTC	CTCTGGAGCT	GCCTGTTTCA	1920
TGAGATGGAG	GAGGTGGGCA	CATTTAAGTT	CAGTTTACTA	ACCTATGGTT	CAGAGTTCTG	1980
ATCATATGGA	AGTTTGGAAA	AGAGAGCTTA	TCACAGGTTT	GTATGCTGGT	GAATGGATAG	2040
TTTTAATTCT	CACTGTCTCA	AAAGAGAATC	AGCTCTCCAG	CAGTTCTAGA	AAAGCTTTGA	2100
CAATCCCCAA	GGGGCAGTGT	TACCTTACTC	CTTCACTGCT	TCTTAGAAGG	TAGAATTAAG	2160
TTTCTGGAAT	TGCACCTACA	TGTTTTCTTA	TTAACATTCA	GAATTGGGAA	TATTAATTTT	2220



5 TCCAGTGAAGT AGTTTTCTGA AATTGGTAAC TTGGAGAGTA AAATAACGTA TTTTGCTTTT 2280  
CAATTTTGTG TTTGTTTACT TTTATGTAAA AATTTGATAT GTGAATTACA CAGTTCTAAT 2340  
AAAACCTCAT GCCTTTTCAT TACATCTAAT TTGAACTCTC AACTTCAGTG CCAGAAGTGC 2400  
TTTAAAGATG CTTTAATGAA AAGTATTAAG AAAATATATA GATTTGTATG TCAGTTTATA 2460  
CTTCAGAAAT CCATATATTT GTCATATTTA TTTTCTTAGA AACCTCCTAA TTGGATAACT 2520  
AGATGGTATT TAAAATGAAT GCCCAAAAAT ATCTTGTACC TTTGTCCAAA AGTTTATCTG 2580  
TTGGAAGCCG CCAGCCATTC ATGTAGAGAG TTTATAAGAA AATAATTTAA AATTGTATGC 2640  
ATTTTATATT ACTATGGTAT CTGTGTACCA TATTTCTAAG TATTCATTAT TAAATTGGTA 2700  
10 CTTCTTAAAA CCATAACCTG GCTTGCCTTT TAGTGTAAAA CACAAAATCC AACATTGTAT 2760  
ATAGAGATTG TTCTTTTATG AAGAAGAGCT GACGTAATTT ATTACCAGTG CATCTGCACA 2820  
AAGACATTAA CATAAGTCTC TGAGCAGTGA TACATTTTCA AACATGAAGA GTGACAACCA 2880  
CCACATTAAA CAACCACGGC AACACTCAGA CTTGGCACTT TCCTACGAAT CCATCCTATA 2940  
TGTGCCTGGT ATCGCCTCTG GCATAACTTA CACGAATCGT CCTCCCTACT TGTCTACGCT 3000  
CCTTCATCAA GCACTTGCCA ACACATTCAC CTCTAACTTG TACAACCTTA CCAACTCACC 3060  
15 ACAACATCTG CAACTCTACC CTATCAACTG CCAACCTAAA GACCCCAAC ACAACACAAC 3120  
CCCCAACAC AAAACCACTA AATCATAACC ACCACACACG CCACACACCA CACACCCACC 3180  
CACACAACCA ACACACCACG ACCAAACACC CCACCACAAA CAAGCTAACA ACCACAAACA 3240  
GACAACACAT CACATACACT CACTACCCCC CCATACTCCC ACCCACCA

20 Seq ID NO: 234 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 27-281

25 1 11 21 31 41 51  
AGCAGGAGGA GAGCTGGCGG GAAGACATGC ACCCCTTGAA GACCCAGAGA GAGGCCGTCT 60  
GTCTACCGCG TAGCAGTTAC ATCAGACTGA GACACTTCCT GTTTACAGGA GACTATAAAA 120  
TTCTTGCCCC GTGCTCATTG GGGGCTGACG CCATTTTAGG CCTCAGCCCA TCTGCACCCA 180  
30 GGCGCTCACT GAAACAGTGT GTTGCTCCAC ACCGCTTGT TTTGCTTGT GGCGCGCTCT 240  
CAGGGTTCCG ACCAATCCAA GAGCCTTGCA GAAAGCATT ACGTGCTTT CTCTTTGGCA 300  
GAGTTTTTCT TTGCTCTGAT CTTGGAGACA TCCCTCTGCC TAGTGGAAAC ATAAGGAATA 360  
CAGAAAGAAT GCAAGGAGAT AGACCAACGT GAGATTCTCC TTCATGCACT CAAGAGAAAG 420  
ATGTTGCAGG AAGAGCTAGT CTTTCAGGCT GGGCTGGTGA CCTGAGAAAG AATGTCCAGC 480  
TTTCTCTCTC CACTTGCCAT ATCAAGAGCC AGGCGTGGAA GACTAAAACA GGAAATGTTT 540  
35 ATAAAAACTG TTCAGCGGTT CGCCAACAAG AAGTGGTAAA GTAGCAAAAA TGGGGATGGA 600  
GATGCCAGGA GGAAAGATGC CAGGGGTAAA GTGGGAAAAT GGGAACCTGA AGCCAGGAGG 660  
TCAAGCCAAG CCAACAGGTG TTCTGTTTTT CATCACAGAA CTAATAAGTG GTGCTGAGGA 720  
CTCAAAACCG GGGAAGCCCA CTCTAGAACC CATGCTGGTC ATCCATATCC CCAAGGCCCT 780  
GGTCAGAAAC CAGCTAAGCA GATGGCTTGG GTCATCAGGA CGTCCATTAC ATCCAAAGGA 840  
40 AGACAGCCTG TGACGTTTCA AAAGCAAAAAG TCCCCTACCA GCCAGTGAAG CTACCTGATT 900  
TCTCAGTATC TTACGCCCAG TGACAGATC TCCCTCAAA ACTTAAAAA AAAAGGGAAA 960  
CATAAACACA TAACAGCAGC AGCAATAATT AAAGATGAGA TGAGAACAAT TAAGAAAAA 1020  
GGAAAGGTCT CCTGTGACTG TTTTATTTTT AGGGAACAG AGAGGAAGAA GAATGATTTT 1080  
TCTTTTGATG ACTCTATATC CAACTCTGAG GTTTGATTAA AGAAATGACC TTGAACCACA 1140  
45 GCAAAGAAAA ATAAAGACA ATTTCCAGTA AGTATGCCAG TTCGAATTAA TGATTACTT 1200  
TTTATTTTTA AACTGAATTC AGCAGAGATT TACATGCATT ACGATGATTA ACATCTGAAA 1260  
TTTGACCTTG AAATAATCTT TACATTGTAA ATTCTTAATG ATCAAAACAA GGTTCCTCAGT 1320  
GATTAAAAAC TATTAGTAAT TAATTATTAA AGGAGAATAA TTGCAAATAC AACATTCCTA 1380  
AAATCTCAAG GCTTTTAAAG CATTTGTACA AATGACTGGA CATTTTTTAA ATTTGAAAAA 1440  
50 AAAAAAAGC CCTCCATCTG ATTCTCATTT TCATTGTGAG TGCAACAACA AAAAAGGTAT 1500  
GCACTTCTCT TCTCATTTTC CACTGTCTCG CAAGCTAGAA ATTCTCACGA CTACCTTTGA 1560  
TCCCATCAA GCAAAGAAA GAAAAGAAAA TTGTCTGTGA CAGATATATG ACATTAAAAA 1620  
ATAATCCC

55 Seq ID NO: 235 Protein sequence:  
Protein Accession #: Eos sequence

60 1 11 21 31 41 51  
MHPLKTQREA VCLPRSSYIR LRHFLFTGDY KIPAPCSFGA DAILGLSPSA PRRSLKQCVA 60  
PHRLVLLVGA LSGFRPIQEP CRKH

65 Seq ID NO: 236 DNA sequence  
Nucleic Acid Accession #: NM\_002075  
Coding sequence: 406..1428

70 1 11 21 31 41 51  
CCACAATAGG GGCAGACCTG TCCATCCTTC TCTGTGGGTC CCCTGTACCT TTCTCCCCCA 60  
ACAGGATCAG ACCCAGAGGC AGCTGGTTGG GGTTCGTCGA GAAGAAGGAT TATCCAGATC 120  
AGTCCTTTCT AATCTCAGCT CCTGCCTGTA CCTCCATA CTCACCAAC CCTCTTCCCC 180  
ACCACCTGGA GCTGAGGAGC ACAGTTTGAG GCCCCCACA CCCCCGCGG GTCGGGGCCA 240  
75 GGCCAGGCCA GGCCAGCTCC TCTGGCAGCA GAGCCTGGGC AGGTGACGGG CGGGCGCGGG 300  
CGTCGCAGCT GAGGGAGTAA GGAGGCTCCC AGGAACCGGA GCTGGAAACC CGGCCGAGGT 360  
CCAGCCAGAG CCCAAGAGCC AGAGTGACCC CTCGACCTGT CAGCCATGGG GGAGATGGAG 420  
CAACTGCGTC AGGAAGCGGA GCAGCTCAAG AAGCAGATTG CAGATGCCAG GAAAGCCTGT 480  
GCTGACGTTA CTGTGGCAGA GCTGGTGTCT GGCCTAGAGG TGGTGGGACG AGTCCAGATG 540  
CGGACGCGGC GGACGTTAAG GGGACACCTG GCCAAGATT ACGCCATGCA CTGGGCCACT 600  
80 GATTCTAAGC TGCTGGTAAG TGCCTCGCAA GATGGGAAGC TGATCGTGTG GGACAGCTAC 660  
ACCACCAACA AGGTGCACGC CATCCCACTG CGCTCCTCCT GGGTCATGAC CTGTGCCTAT 720  
GCCCCATCAG GGAACCTTGT GGCATGTGGG GGGCTGGACA ACATGTGTTT CATCTACAAC 780  
CTCAAATCCC GTGAGGGCAA TGTCAAGGTC AGCCGGGAGC TTTCTGCTCA CACAGGTTAT 840  
85 CTCTCCTGCT GCCGCTTCTT GAGTGAACA AATATTGTGA CCAGCTCGGG GGACACCACG 900  
TGTGCCTTGT GGGACATTGA GACTGGGCAG CAGAAGACTG TATTTGTGGG ACACACGGGT 960  
GACTGCATGA GCCTGGCTGT GTCTCCTGAC TTCAATCTCT TCATTTCCGG GGCCTGTGAT 1020  
GCCAGTGCCA AGCTCTGGGA TGTGCGAGAG GGGACCTGCC GTCAGACTTT CACTGGCCAC 1080



GAGTCGGACA TCAACGCCAT CTGTTTCTTC CCCAATGGAG AGGCCATCTG CACGGGCTCG 1140  
 GATGACGCTT CCTGCCGCTT GTTTGACCTG CGGGCAGACC AGGAGCTGAT CTGCTTCTCC 1200  
 CACGAGAGCA TCATCTGCGG CATCACGTCC GTGGCCTTCT CCCTCAGTGG CCGCCTACTA 1260  
 TTCGCTGGCT ACGACGACTT CAACTGCAAT GTCTGGGACT CCATGAAGTC TGAGCGTGTG 1320  
 5 GGCATCCTCT CTGGCCACGA TAACAGGGTG AGCTGCCTGG GAGTCACAGC TGACGGGATG 1380  
 GCTGTGGCCA CAGGTTCTCT GGACAGCTTC CTCAAATCT GGAAGTGAAG AGGCTGGAGA 1440  
 AAGGGAAGTG GAAGGCAGTG AACACACTCA GCAGCCCCCT GCCCGACCCC ATCTCATTTCA 1500  
 GGTGTTCTCT TCTATATTCC GGGTGCCATT CCCACTAAGC TTTCTCCTTT GAGGGCAGTG 1560  
 10 GGGAGCATGG GACTGTGCCT TTGGGAGGCA GCATCAGGGA CACAGGGGCA AAGAAGTGGC 1620  
 CCATCTCCTC CCATGGCCTT CCCTCCCCAC AGTCTCACA GCCTCTCCCT TAATGAGCAA 1680  
 GGACAACCTG CCCCTCCCCA GCCCTTTGCA GGCCAGCAG ACTTGAGTCT GAGGCCCCAG 1740  
 GCCCTAGGAT TCCTCCCCCA GAGCCACTAC CTTTGTCCAG GCCTGGGTGG TATAGGGCGT 1800  
 TTGGCCCTGT GACTATGGCT CTGGCACCAC TAGGGTCTCT GCCCTCTTCT TATTCATGCT 1860  
 15 TTCTCCTTTT TCTACCTTTT TTTCTCTCCT AAGACACCTG CAATAAAGTG TAGCACCTG 1920  
 GT

Seq ID NO: 237 Protein sequence:  
 Protein Accession #: NP\_002066

1 11 21 31 41 51  
 | | | | |  
 MGEMEQLRQE AEQLKKQIAD ARKACADVTL AELVSGLEV GRVQMRTRRT LRHGLAKIYA 60  
 MHWATDSKLL VSASQDGKLI VWDSYTTNKV HAIPLRSSWV MTCAYAPSGN FVACGGLDNM 120  
 25 CSIYNLKSRE GNVKVSRELS AHTGYLSCCR FLDDNNIVTS SGDTTCALWD IETGQKQTVF 180  
 VGHTGDCMSL AVSPDFNLFI SGACDASAKL WDVREGTCRQ TFGHESDIN AICFFPNGEA 240  
 ICTGSDDASC RLFDLRADQE LICFSHESII CGITSVAFSL SGRLLFAGYD DFNCNVWDSM 300  
 KSERVGILSG HDNRVSLGVL TADGMAVATG SWDSFLKIWN

Seq ID NO: 238 DNA sequence  
 Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
 | | | | |  
 TCCCAATGTG TNGAACCTAC CATAAATTCT TTTCTTACNG GACAATCTTA TNCTAANCAA 60  
 35 TACCATTTCG TTTTAAGGCA GATAATCCTC CAAGTTTTCT AATGATATCT GAAACTATTA 120  
 ACTGATTCTG TGAATTATGA AATCTGAAAA GGAATTGGAA GTTGCTAAAA ATCTATCATT 180  
 TGCATTGACC AGTGTGAAGC ACAGTGGAAAT GAGAATGCGT GCCCTGACAC CAAAGAAAAA 240  
 TAAGTGAAGT GAAAGCTGAA GAATCACCGG CTTCAGTGAC ATGGAACCCA GTGATTTGAT 300  
 TTTTGACGAG TATCGGGTGA CTTTGAGGTG GTCAAGAAAC CACACTTTAA GAACAATGTC 360  
 40 CAAAAAGGGG AAAAAAAGA GCAACCAAAG AAAAAAATC CATAAAATTG CACAGAAGAA 420  
 AAGAAAGAAA AATAAAATAC ACAATATGGA CGATGGAGAA AAACAGTTAC ATTTCTTTAT 480  
 GGATCAAGAA GTTTGTGTAC ACATAATCTC ATTTTGAGAT ATATAACTAT TTTTGTCTTT 540  
 CAGAAGTGAA TCAAAATATT TCAAAATGCT GTCTTATGAA ACTACAATAT TCTCACAGAT 600  
 TAGAAAAGTT TTTCTGTAAA AGTCAGATAG TAAATATTTT AGGTTTTGCA GTGTCTTTTG 660  
 45 CAACTACTCA ACTTTCCTAC TGTCAGACAA GAGTAGCTGT GGTACTGTGC AAATAAATTG 720  
 CTTGTGTTCC AATAAAGCTT CATTTACAAA AACATGCCAT GGGCCATATT TGGCCTGTAC 780  
 ACTGTTGTTT GCCAAGTCCT AATATAGTTT CTTAGCAAGT ATTGTGAGCT ATTTGAGGAA 840  
 GACATGAAAG TTCATTGGGT TGCTAAAAAG TATGTAGAAA TTCAAAGGAA AATTAAATT 900  
 50 TAGGCTAAGT TATAATACAC TGTTTTAACA ATTGTAAAT GTAAGAGAAA TTTACAAATA 960  
 AAAATCCCAA ATAAAA

Seq ID NO: 239 DNA sequence  
 Nucleic Acid Accession #: NM\_001786.1  
 Coding sequence: 130-1023

1 11 21 31 41 51  
 | | | | |  
 GGGGGGGGGG GGCACCTGGC TTCAAAGCTG GCTCTTGGAA ATTGAGCGGA GAGCGACGCG 60  
 60 GTTGTGTAG CTGCCGCTGC GGCCGCCGCG GAATAATAAG CCGGGATCTA CCATACCCAT 120  
 TGACTAATA TGGAAGATTA TACCAAAATA GAGAAAATTG GAGAAGGTAC CTATGGAGTT 180  
 GTGTATAAGG GTAGACACAA AACTACAGGT CAAGTGGTAG CCATGAAAAA AATCAGACTA 240  
 GAAAGTGAAG AGGAAGGGGT TCCTAGTACT GCAATTCCGG AAATTTCTCT ATTAAGGAA 300  
 CTTGTCATC CAAATATAGT CAGTCTTCAG GATGTGCTTA TGCAGGATTC CAGGTTATAT 360  
 65 CTCATCTTTG AGTTTCTTTC CATGGATCTG AAGAAATACT TGGATTCTAT CCCTCCTGGT 420  
 CAGTACATGG ATTCTTCACT TGTTAAGAGT TATTTATACC AAATCCTACA GGGGATTGTG 480  
 TTTTGTCACT CTAGAAGAGT TCTTCACAGA GACTTAAAC CTCAAATCT CTGATTGAT 540  
 GACAAAGGAA CAATTAAGT GGCTGATTTT GCCTTGCCA GAGCTTTTGG AATACCTATC 600  
 AGAGTATATA CACATGAGGT AGTAACACTC TGGTACAGAT CTCCAGAAGT ATTGCTGGGG 660  
 70 TCAGCTCGTT ACTCAACTCC AGTTGACATT TGGAGTATAG GCACCATATT TGCTGAACATA 720  
 GCAACTAAGA AACCCTTTT CCATGGGGAT TCAGAAATTG ATCAACTCTT CAGGATTTTC 780  
 AGAGCTTTGG GCACTCCCAA TAATGAAGTG TGGCCAGAAG TGAATCTTT ACAGGACTAT 840  
 AAGAATACAT TTCCCAAATG GAAACCAGGA AGCCTAGCAT CCCATGTCAA AACTTGGAT 900  
 GAAAAATGGCT TGGATTTGCT CTCGAAAATG TTAATCTATG ATCCAGCCAA ACGAATTTCT 960  
 75 GGCAAAATGG CACTGAATCA TCCATATTTT AATGATTTGG ACAATCAGAT TAAGAAGATG 1020  
 TAGCTTTCTG AAAAAAGTT TCCATATGTT ATGTCAACAG ATAGTTGTGT TTTTATTTGT 1080  
 AACTCTTGTG TATTTTGTG TTATATATAT TTCTTTGTTA TCAAACCTCA GCTGTACTTC 1140  
 GTCTTCTAAT TTCAAAAATA TAACCTAAAA ATGTAAATAT TCTATATGAA TTTAAATATA 1200  
 ATCTGTAAAA TGTGAAAAAA AAAAAAATA AAAAA

Seq ID NO: 240 Protein sequence:  
 Protein Accession #: NP\_001777.1

1 11 21 31 41 51  
 | | | | |  
 MEDYTKIEKI GEGTYGVVYK GRHKTGQVV AMKKIRLESE EEGVPSTAIR EISLLKELRH 60  
 85 PNIVSLQDVL MQDSRLYLIF EFLSMDLKKY LDSIPPQYMD DSSLVKSYLY QILQGIVFCH 120



SRRVLHRDLK PQNLLIDDKG TIKLADFLGA RAFGIPIRVY THEVVTWLWYR SPEVLLGSAR 180  
 YSTPVDIWSI GTIFAEELATK KPLFHDSEI DQLFRIFRAL GTPNNEVWPE VESLQDYKNT 240  
 FPKWKPGSLA SHVKNLDENG LDLLSKMLIY DPAKRISGKM ALNHPYFNDL DNQIKKM

5

Seq ID NO: 241 DNA sequence  
 Nucleic Acid Accession #: NM\_033379.1  
 Coding sequence: 132-854

10

1	11	21	31	41	51	
CGCCCGCGCG	CGGGCTCAAC	TTTGTAGAGC	GAGGGGCCAA	CTTGGCAGAG	CGCGCGGCCA	60
GCTTTGCAGA	GAGCGCCCTC	CAGGGACTAT	GCGTGCGGGG	ACACGGGATC	TACCCATACC	120
ATTGACTAAC	TATGGAAGAT	TATACCAAAA	TAGAGAAAAT	TGGAGAAGGT	ACCTATGGAG	180
TTGTGTATAA	GGGTAGACAC	AAAACCTACG	GTCAAGTGGT	AGCCATGAAA	AAAATCAGAC	240
TAGAAAAGTGA	AGAGGAAGGG	GTTCCTAGTA	CTGCAATTCG	GGAAATTTCT	CTATTAAAGG	300
AACTTCGTCA	TCCAAATATA	GTCAGTCTTC	AGGATGTGCT	TATGCAGGAT	TCCAGGTTAT	360
ATCTCATCTT	TGAGTTTCTT	TCCATGGATC	TGAAGAAATA	CTTGGATTCT	ATCCCTCCTG	420
GTCAGTACAT	GGATTCTTCA	CTTGTAAAGG	TAGTAACACT	CTGGTACAGA	TCTCCAGAAG	480
TATTGCTGGG	GTCAGCTCGT	TACTCAACTC	CAGTTGACAT	TTGGAGTATA	GGCACCATAT	540
TTGCTGAAC	AGCAACTAAG	AAACCACTTT	TCCATGGGGA	TTCAGAAATT	GATCAACTCT	600
TCAGGATTTT	CAGAGCTTTG	GGCACTCCCA	ATAATGAAGT	GTGGCCAGAA	GTGGAATCTT	660
TACAGGACTA	TAAGAATACA	TTTCCCAAAAT	GGAAACCAGG	AAGCCTAGCA	TCCCATGTCA	720
AAAACCTTGA	TGAAAATGGC	TTGGATTTCG	TCTCGAAAAT	GTTAATCTAT	GATCCAGCCA	780
AACGAATTTT	TGGCAAAATG	GCACTGAATC	ATCCATATTT	TAATGATTG	GACAATCAGA	840
TTAAGAAGAT	GTAGCTTTCT	GACAAAAAGT	TTCCATATGT	TATGTCAACA	GATAGTTGTG	900
TTTTTATTGT	TAACCTTTGT	CTATTTTGT	CTTATATATA	TTTCTTTGTT	ATCAAACTTC	960
AGCTGTACTT	CGTCTTCTAA	TTTCAAAAAT	ATAACTTAAA	AATGTAAATA	TTCTATATGA	1020
ATTTAAATAT	AATTCTGTAA	ATGTGAAAAA	AAAAAAAAAA	AAAAAA		

30

Seq ID NO: 242 Protein sequence:  
 Protein Accession #: NP\_203698.1

35

1	11	21	31	41	51	
MEDYTKIEKI	GEGTYGVVYK	GRHKTTGQVV	AMKKIRLESE	EEGVPSTAIR	EISLLKELRH	60
PNIVSLQDVL	MQDSRLYLIF	EFLSMDLKKY	LDSIPPGQYM	DSSLVKVVTL	WYRSPEVLLG	120
SARYSTPVDI	WSIGTIFAEI	ATKKPLFHD	SEIDQLFRIF	RALGTPNNEV	WPEVESLQDY	180
KNTFPKWKPG	SLASHVKNLD	ENGLDLSKLM	LIYDPAKRIS	GKMLNHPYF	NDLDNQIKKM	

40

Seq ID NO: 243 DNA sequence  
 Nucleic Acid Accession #: AF101051.1  
 Coding sequence: 221-856

45

1	11	21	31	41	51	
GAGCAACCTC	AGCTTCTAGT	ATCCAGACTC	CAGCGCCGCC	CCGGGCGCGG	ACCCCAACCC	60
CGACCCAGAG	CTTCTCCAGC	GGCGGCGCAG	CGAGCAGGGC	TCCCCGCCTT	AACTTCCTCC	120
GCGGGGCCCA	GCCACCTTCG	GGAGTCCGGG	TTGCCACCT	GCAAACCTC	CGCCTTCTGC	180
ACCTGCCACC	CCTGAGCCAG	CGCGGGCGCC	CGAGCGAGTC	ATGGCCAACG	CGGGGCTGCA	240
GCTGTTGGGC	TTCACTCTCG	CCTTCCTGGG	ATGGATCGGC	GCCATCGTCA	GCACTGCCCT	300
GCCCCAGTGG	AGGATTTACT	CCTATGCCGG	CGACAACATC	GTGACCGCCC	AGGCCATGTA	360
CGAGGGGCTG	TGGATGTCCT	GCGTGTGCGA	GAGCACCGGG	CAGATCCAGT	GCAAAGTCTT	420
TGACTCCTTG	CTGAATCTGA	GCAGCACATT	GCAAGCAACC	CGTGCCTTGA	TGGTGGTTGG	480
CATCCTCCTG	GGAGTGATAG	CAATCTTTGT	GGCCACCGTT	GGCATGAAGT	GTATGAAGTG	540
CTTGGAAGAC	GATGAGGTGC	AGAAGATGAG	GATGGCTGTC	ATTGGGGGTG	CGATATTTCT	600
TCTTGACAGT	CTGGCTATTT	TAGTTGCCAC	AGCATGGTAT	GGCAATAGAA	TCGTTCAAGA	660
ATTCTATGAC	CCTATGACCC	CAGTCAATGC	CAGGTACGAA	TTTGGTCAGG	CTCTCTTCAC	720
TGGCTGGGCT	GCTGCTTCTC	TCTGCCTTCT	GGGAGGTGCC	CTACTTTGCT	GTTCTGTGCC	780
CCGAAAAACA	CAACCTTACC	GCCCTATCCA	AAACCTGCAC	CTTCCAGCGG		840
GAAAGACTAC	GTGTGACACA	GAGGCAAAAG	GAGAAAATCA	TGTTGAAACA	AACCGAAAAT	900
GGACATTGAG	ATACTATCAT	TAACATTAGG	ACCTTAGAAT	TTTGGGTATT	GTAATCTGAA	960
GTATGGTATT	ACAAAACAAA	CAAAACAAACA	AAAAACCCAT	GTGTTAAAT	ACTCAGTGCT	1020
AAACATGGCT	TAATCTTATT	TTATCTTCTT	TCCTCAATAT	AGGAGGGAAG	ATTTTACCAT	1080
TTGTATTACT	GCTTCCCATT	GAGTAATCAT	ACTCAAATGG	GGGAAGGGGT	GCTCCTTAAA	1140
TATATATAGA	TATGTATATA	TACATGTAT	TCTATAAAA	ATAGACAGTA	AAATACTATT	1200
CTCATTATGT	TGATACTAGC	ATACTTAAAA	TATCTCTAAA	ATAGGTAAAT	GTATTTAATT	1260
CCATATTGAT	GAAGATGTTT	ATTGGTATAT	TTTCTTTTTC	GTCTTATAT	ACATATGTAA	1320
CAGTCAAATA	TCATTTACTC	TTCTTCATTA	GCTTTGGGTG	CCTTTGCCAC	AAGACCTAGC	1380
CTAATTTACC	AAGGATGAAT	TCTTTCAATT	CTTCATGCGT	GCCCTTTTCA	TATACTTATT	1440
TTATTTTSTA	CCATAATCTT	ATAGCACTTG	CATCGTTATT	AAGCCCTTAT	TTGTTTGTG	1500
TTTCATTGGT	CTCTATCTCC	TGAATCTAAC	ACATTTTATA	GCCTACATTT	TAGTTTCTAA	1560
AGCCAAGAAG	AATTTATTAC	AAATCAGAAC	TTTGGAGGCA	AATCTTTCTG	CATGACCAAA	1620
GTGATAAATT	CCTGTTGACC	TTCCACACACA	ATCCCTGTAC	TCTGACCCAT	AGCACTCTTG	1680
TTTGCTTTGA	AAATATTGT	CCAAATGAGT	AGCTGCATGC	TGTTCCCCCA	GGTGTGTGAA	1740
CACAACTTTA	TTGATTGAAT	TTTAAAGCTA	CTTATTCTATA	GTTTTATATC	CCCCTAAACT	1800
ACCTTTTGTG	TCCCCATTCC	TTAATTGTAT	TGTTTTCCCA	AGTGTAATTA	TCATGCGTTT	1860
TATATCTTCC	TAATAAGGTG	TGGTCTGTTT	GTCTGAACAA	AGTGCTAGAC	TTTCTGGAGT	1920
GATAATCTGG	TGACAAATAT	TCTCTCTGTA	GCTGTAAGCA	AGTCACTTAA	TCTTTCTACC	1980
TCTTTTTTCT	ATCTGCCAAA	TTGAGATAAT	GATACTTAAC	CAGTTAGAAG	AGGTAGTGTG	2040
AATATTAATT	AGTTTATATT	ACTCTCATTC	TTTGAACATG	AACTATGCCT	ATGTAGTGTC	2100
TTTATTTGCT	CAGCTGGCTG	AGACACTGAA	GAAGTCACTG	AACAAAACCT	ACACACGTAC	2160
CTTCATGTGA	TTCACTGCCT	TCCTCTCTCT	ACCAGTCTAT	TTCCACTGAA	CAAAACCTAC	2220
ACACATACCT	TCATGTGGTT	CAGTGCCTTC	CTCTCTCTAC	CAGTCTATTT	CCACTGAACA	2280
AAACCTACGC	ACATACCTTC	ATGTGGCTCA	GTGCCTTCCT	CTCTCTACCA	GTCTATTTCC	2340
ATTCTTTTCT	CTGTGTCTGA	CATGTTTGTG	CTCTGTTCCT	TTTTAACAAC	TGCTCTTACT	2400
TTTCCAGTCT	GTACAGAATG	CTATTTCACT	TGAGCAAGAT	GATGTATGGA	AAGGGTGTG	2460

85



5 GCACTGGTGT CTGGAGACCT GGATTTGAGT CTTGGTGCTA TCAATCACCG TCTGTGTTTG 2520  
AGCAAGGCAT TTGGCTGCTG TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAATTC 2580  
CTGATCTTCC CACCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640  
GTGGTTTTGT AATTTGAAAA GTGCTATACT AAGGGAAAAG ATTGAGGAAT TAACTGCATA 2700  
CGTTTTGGTG TTGCTTTTCA AATGTTTGAA AATAAAAAAA TGTTAAGAAA TGGGTTTCTT 2760  
GCCTTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAAACGAAT 2820  
AAGATTCTGA GGAAGTCTTA TCTTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880  
ACAGATGTAA TGGGAAGAAA TAAAAGCCTA CGTGTGTTGA AATCCAACAG CAAGGGAGAT 2940  
TTTGAATCA TAATAACTCA TAAGGTGCTA TCTGTTTCACT GATGCCCTCA GAGCTCTTGC 3000  
10 TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGGTACT TCATAATAAAA 3060  
CTACACAAGG AAAGTCAGCC ACCGTGTCTT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120  
TGCTTCCAA ACCTGAGAAT ATATGCTTTT GGAAGTTAAA ATTTAAATGG CTTTTGCCAC 3180  
ATACATAGAT CTTTCATGATG TGTGAGTGTA ATTCATGTG GATATCAGTT ACCAAACATT 3240  
ACAAAAAAT TTTATGGCCC AAAATGACCA ACGAAATTGT TACAATAGAA TTTATCCAAT 3300  
15 TTGATCTTT TTATATTCTT CTACCACACC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360  
TTATAATGGG AATTTGTATA AAGCATTACT CTTTTTCAAT AAATTGTTTT TTAATTTAAA 3420  
AAAAGGAAAA AAAAAAAAAA AAA

20 Seq ID NO: 244 Protein sequence:  
Protein Accession #: AAD16433.1

25 1 11 21 31 41 51  
| | | | | |  
MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMY EGL WMSCVSQSTG 60  
QIQCKVFDL LNLSSLTQAT RALMVVGILL GVIAIFVATV GMKCMKLED DEVQKMRMAV 120  
IGGAIFLLAG LAILVATAWY GNRIVQEFYD PMTPVNARYE FGQALFTGWA AASLCLLGA 180  
LLCCSCPRKT TSYPTPRPYP KPAPSSGKDY V

30 Seq ID NO: 245 DNA sequence  
Nucleic Acid Accession #: CAT cluster

35 1 11 21 31 41 51  
| | | | | |  
TTTTTTTTTT TTTTTTTTTT TTTTCAAGG AGAGCACAAAG GAACTTTATT AATGACTTTC 60  
TTAATGGTTA AATGCTGTTT ACCAAGTGAC CCAGAGGCAG CGTGGTTTAG TGGTTTCAAC 120  
AGCATGGTCC CGAGAGTCTG ACAAACCTCA GTTCAAATCC TTCTTTTGTC TTCACTTAGT 180  
TTTCTTCCT GAGATTTAGT TTCTTCATCG TTAACAATGA GGATATTAAT ATGTTTCACA 240  
CAGTTGTTAT GAAGAATGCA TATATTAGAA TGCCTGTAGT CTCAGCTACT CAGGAGGCTA 300  
40 AGGTGGGGAG GTCGCTCAAG CCCAGGAATT CAAAGCTGCA ATGCATTATG ATTACAGCTG 360  
TTAATAGCCA CTGCACTTCA GCCTGGGCAA TGTAAGTAAGA TCCCATCTCT GGCTCGGAGG 420  
GTCCTACGCC CACGGAGTCT CGCTGATTGC TAGCACAGCA GTCTGAGATC AAAGTGCA

45 Seq ID NO: 246 DNA sequence  
Nucleic Acid Accession #: XM\_058553.2  
Coding sequence: 897-1400

50 1 11 21 31 41 51  
| | | | | |  
AATTTTCAGA AGTTTCGTAT GGGGATGGTT TTATATAAAT TCAGGTTTTT CCCACAATAA 60  
TAAATGTATT TAGTCTCAGT GCTCAATAGA AGAGATTTCT AATAGAAAAG GATTCAAAC 120  
GTGAAACCAT TTCTCTTTTA ATGTTTCACA TTCTTTTAC AGATTGTGTC TCTTGTGACT 180  
CTGTTATCCA TAATATGGAC AGTTCTTGAG TCCTAACATG GAGAGGTTTT CCCTTAGTGC 240  
ATAGAGGGAA TGAGATTAA TTGGAGAAGC TTAAAGTATT GCCACTTAG CACTGAAGAT 300  
55 TGGGATGAGA GGAGGTGAAA CCTCACTAGA AAAAGGGACA ATGTTAGTGT GGCCCTTCCT 360  
GATCATGTTT AAGAAAAGTC ATGAAAATGG TGAAGTAGTG TTTCCAAGCA TATTGGAAGG 420  
GTTGAGTGTA TACTGTCTGT CAAAGACTTC CAGCATTTC AGGTCTAGA GAGGAACAAG 480  
ACTGGTAACC TGCCTATCTG TATTTTTAAG AACCCAGGAG GAAAGCTTTA TAATAGAACA 540  
TTATTTCTGT GTTTATGTAT AAGGGGTTTT TTGTTTTTTT AAAGACAGGA TCTCACTCCA 600  
60 TTGTCCAGGC CAAGTGCAAT GGCACGAACC TCATAGCTCC TGGACTTAAG TGATCTGCCT 660  
GCCTTTGCCT CCTGAGTAGC TGGGACTACA GGCATGAGCC CCCATGCCTG GCTAAGTTTG 720  
TTTTTTTGTG TGTTTGTGTT TTTGTTTTTG GGGGGGGTTG TTTTGTTTTT TGTAAGAGACG 780  
TAGTCTTGCT TTGTTGCCAG GCTAGTCTCA AACTCTGGC TTCAAGTGAT CCTCCTGCCT 840  
CAGCCTCCCA GAGTGCTAGG ATTACAGCAC TTGGATTGAG CTTCTTCATT TCCAACATGG 900  
65 AAGAACTTA CACCGACTCC CTGGACCTAT AGAAGCTATT GCAATGCCCC TATGACAAAA 960  
ACCATCAAAT CAGGGCTTGC AGGTTTCCTT ATCATCTTAT CAAGTGCAGA AAGAATCATC 1020  
CTGATGTTGC AAGCAAATTG GCTACTTGTC CCTTCAATGC TCGCCACCAG GTTCTCGAG 1080  
CTGAAATTAG TCATCATATC TCAAGCTGTG ATGACAGAAG TTGTATTGAG CAAGATGTTG 1140  
TCAACCAAAC CAGGAGCCTT AGACAAGAGA CTCTGGCTGA GAGCACTGG CAGTGCCCTC 1200  
70 CTTGCGATGA AGACTGGGAT AAAGATTGTT GGGAGCAGAC CAGCACCCCA TTTGTCTGGG 1260  
GCACAACTCA CTACTCTGAC AACAACAGCA CTGCGAGCAA CATAGTTACA GAACATAAGA 1320  
ATAACCTGGC TTCAGGCATG CGAGTTCCCA AATCTCTGCC GTATGTTCTG CCATGGAAAA 1380  
ACAAATGGAAA TGCACAGTAA CTGAATACCT ATCTCATCAA ATGCCAGACC CTAGAAGACT 1440  
75 GTTGCTTCTT CTTCTACCAG TGGGTTCTCA TTTTCTCTCT AATCTAATTA TAGAATGGTA 1500  
AACTCCCTGT GACTTTCCAA ACTGACAAGC ACACTTTTT CCTCCCCCT TGAATCCTCA 1560  
TTTAATGCAA GAACCTCAT ACTCAGAAGC TTCCAAATAA ACCTTTGATA CAGATTG

80 Seq ID NO: 247 Protein sequence:  
Protein Accession #: XP\_058553.1

85 1 11 21 31 41 51  
| | | | | |  
MEETYDSDL PEKLLQCPYD KNHQIRACRF PYHLIKCRKN HPDVASKLAT CPFNARHQVP 60  
RAEISHHIS CDDRSCIEQD VVNQTRSLRQ ETLAESTWQC PPCDEDWDKD LWEQTSTPFV 120  
WGTTHYSNN SPASNIVTEH KNNLASGMRV PKSLPYVLPW KNNGNAQ



Seq ID NO: 248 DNA sequence  
Nucleic Acid Accession #: NM\_003392  
Coding sequence: 758..1855

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80

1	11	21	31	41	51	
TTAAGGAAAT	CCGGGCTGCT	CTTCCCCATC	TGGAAGTGGC	TTTCCCCACA	TCGGCTCGTA	60
AACTGATTAT	GAAACATACG	ATGTTAATTC	GGAGCTGCAT	TTCCCAGCTG	GGCACTCTCG	120
CGCGCTGGTC	CCCGGGCCT	CGCCCCCAC	CCCCTGCCCT	TCCCTCCCGC	GTCCTGCCCC	180
CATCCTCCAC	CCCCCGCGCT	GGCCACCCCG	CCTCCTTGGC	AGCCTCTGGC	GGCAGCGCGC	240
TCCACTCGCC	TCCCGTGCTC	CTCTCGCCCA	TGGAATTAAT	TCTGGCTCCA	CTTGTTGCTC	300
GGCCCAGGTT	GGGGAGAGGA	CGGAGGGTGG	CCGCAGCGGG	TTCCTGAGTG	AATTACCCAG	360
GAGGGACTGA	GCACAGCACC	AACTAGAGAG	GGGTCAAGGG	GTGCGGGACT	CGAGCGAGCA	420
GGAAGGAGGC	AGCGCCTGGC	ACCAGGGCTT	TGACTCAACA	GAATTGAGAC	ACGTTTGTAA	480
TCGCTGGCGT	GCCCCGCGCA	CAGGATCCCA	GCGAAAATCA	GATTTCTCTG	TGAGGTTGCG	540
TGGGTGGATT	AATTTGGAAA	AAGAACTGCT	CTATATCTTG	CCATCAAAAA	ACTCACGGAG	600
GAGAAGCGCA	GTCAATCAAC	AGTAACTTA	AGAGACCCCC	GATGCTCCCC	TGGTTTAACT	660
TGTATGCTTG	AAAAATTATCT	GAGAGGGAAT	AAACATCTTT	TCCTTCTTCC	CTCTCCAGAA	720
GTCCATTGGA	ATATTAAGCC	CAGGAGTTGC	TTTGGGGATG	GCTGGAAGTG	CAATGTCTTC	780
CAAGTTCTTC	CTAGTGGCTT	TGGCCATATT	TTTCTCCTTC	GCCCAGGTTG	TAATTGAAGC	840
CAATTCTTGG	TGGTCGCTAG	GTATGAATAA	CCCTGTTTCT	ATGTCAGAAG	TATATATTAT	900
AGGAGCACAG	CCTCTCTGCA	GCCAACTGGC	AGGACTTTCT	CAAGGACAGA	AGAACTGTG	960
CCACTTGTAT	CAGGACCACA	TGCAGTACAT	CGGAGAAGGC	GCGAAGACAG	GCATCAAAGA	1020
ATGCCAGTAT	CAATTCCGAC	ATCGACGGTG	GAACTGCAGC	ACTGTGGATA	ACACCTCTGT	1080
TTTTTGGCAGG	GTGATGCAGA	TAGGCAGCCG	CGAGACGGCC	TTACATACG	CCGTGAGCGC	1140
AGCAGGGGTG	GTGAACGCCA	TGAGCCGGGC	GTGCCGCGAG	GGCGAGCTGT	CCACCTGCGG	1200
CTGCAGCCGC	GCCGCGCGCC	CCAAGGACCT	GCCGCGGGAC	TGGCTCTGGG	GCGGCTGCGG	1260
CGACAACATC	GACTATGGCT	ACCGCTTTGC	CAAGGAGTTC	GTGGACGCCC	GCGAGCGGGA	1320
GCGCATCCAC	GCCAAGGGCT	CCTACGAGAG	TGCTCGCATC	CTCATGAACC	TGCACAACAA	1380
CGAGGCCGGC	CGCAGGACGG	TGTACAACTT	GGCTGATGTG	GCCTGCAAGT	GCCATGGGGT	1440
GTCCGGCTCA	TGTAGCCTGA	AGACATGCTG	GCTGCAGCTG	GCAGACTTCC	GCAAGGTGGG	1500
TGATGCCCTG	AAGGAGAAGT	ACGACAGCGC	GGCGGCCATG	CGGCTCAACA	GCCGGGGCAA	1560
GTTGGTACAG	GTCAACAGCC	GCTTCAACTC	GCCCCACCAC	CAAGACCTGG	TCTACATCGA	1620
CCCCAGCCCT	GACTACTGCG	TGCGCAATGA	GAGCACCGGC	TCGCTGGGCA	CGCAGGGCCG	1680
CCTGTGCAAC	AAGACGTCGG	AGGGCATGGA	TGGCTGCGAG	CTCATGTGCT	GCGGCCGTGG	1740
GTACGACCAG	TTCAAGACCG	TGCAGACGGA	GCGCTGCCAC	TGCAAGTTCC	ACTGGTGCTG	1800
CTACGTCAAG	TGCAAGAAGT	GCACGGAGAT	CGTGGAACAG	TTTGTGTGCA	AGTAGTGGGT	1860
GCCACCCAGC	ACTCAGCCCC	GCTCCCAGGA	CCCGCTTATT	TATAGAAAGT	ACAGTGATTC	1920
TGGTTTTTGG	TTTTTTAGAA	TATTTTTTAT	TTTTTCCCAA	GAATTGCAAC	CGGAACCATT	1980
TTTTTTCTTG	TTACCATCTA	AGAACTCTGT	GGTTTATTAT	TAATATTATA	ATTATTATTT	2040
GGCAATAATG	GGGGTGGGAA	CCACGAAAAA	TATTTATTTT	GTGGATCTTT	GAAAAGGTAA	2100
TACAAGACTT	CTTTTGGATA	GTATAGAAAT	AAGGGGGAAA	TAACACATAC	CCTAACTTAG	2160
CTGTGTGGGA	CATGGTACAC	ATCCAGAAGG	TAAAGAAATA	CATTTTCTTT	TTCTCAAATA	2220
TGCCATCATA	TGGGATGGGT	AGGTTCCAGT	TGAAAGAGGG	TGGTAGAAAT	CTATTCACAA	2280
TTAGCTTCT	ATGACCAAAA	TGAGTTGTAA	ATTCTCTGGT	GCAAGATAAA	AGGTCTTGGG	2340
AAAACAAAAC	AAAACAAAAC	AAACCTCCCT	TCCCCAGCAG	GGCTGCTAGC	TTGCTTTCTG	2400
CATTTTCAAA	ATGATAATTT	ACAATGGAAG	GACAAGAATG	TCATATTCTC	AAGGAAAAAA	2460
GGTATATCAC	ATGTCTCATT	CTCCTCAAAT	ATTCCATTTG	CAGACAGACC	GTCAATTCTT	2520
AATAGCTCAT	GAAATTTGGG	CAGCAGGGAG	GAAAGTCCCC	AGAAATTAAA	AAATTTAAAA	2580
CTCTTATGTC	AAGATGTTGA	TTTGAAGCTG	TTATAAGAAAT	TGGGATTCCA	GATTTGTAAA	2640
AAGACCCCCA	ATGATTCTGG	ACACTAGATT	TTTTGTTTGG	GGAGGTTGGC	TTGAACATAA	2700
ATGAAATATC	CTGTATTTTC	TTAGGGATAC	TTGGTTAGTA	AATTATAATA	GTAGAAATAA	2760
TACATGAATC	CCATTACACG	GTTTCTCAGC	CCAAGCAACA	AGGTAATTGC	GTGCCATTCA	2820
GCACTGCACC	AGAGCAGACA	ACCTATTTGA	GGAAAAACAG	TGAAATCCAC	CTTCTCTTTC	2880
AACTAGAGCC	CTCTCTGATT	CCTCCGTGTT	GTGATGTGAT	GCTGGCCACG	TTTCCAAACG	2940
GCAGCTCCAC	TGGGTCCCCCT	TTGGTTGTAG	GACAGGAAAT	GAAACATTAG	GAGCTCTGCT	3000
TGGAAAACAG	TTCACTACTT	AGGGATTTTT	GTTTCCTAAA	ACTTTTATTT	TGAGGAGCAG	3060
TAGTTTTCTA	TGTTTTAATG	ACAGAACTTG	GCTAATGGAA	TTACACAGAG	TGTTGCAGCG	3120
TATCACTGTT	ATGATCCTGT	GTTTAGATTA	TCCACTCATG	CTTCTCCTAT	TGTACTGCAG	3180
GTGTACCTTA	AAACTGTTCC	CAGTGTACTT	GAACAGTTGC	ATTTATAAGG	GGGGAAATGT	3240
GGTTTAAATG	TGCCTGATAT	CTCAAAGTCT	TTTGTACATA	ACATATATAT	ATATATACAT	3300
ATATATAAAT	ATAAATATAA	ATATATCTCA	TTGCAGCCAG	TGATTTAGAT	TTACAGCTTA	3360
CTCTGGGGTT	ATCTCTCTGT	CTAGAGCATT	GTGTCTCTTC	ACTGCAGTCC	AGTTGGGATT	3420
ATTCAAAAAG	TTTTTTGAGT	CTTGAGCTTG	GGCTGTGGCC	CCGCTGTGAT	CATACCCTGA	3480
GCACGACGAA	GCAACCTCGT	TTCTGAGGAA	GAAGCTTGAG	TTCTGACTCA	CTGAAATGCG	3540
TGTTGGGTTG	AAGATATCTT	TTTTTCTTTT	CTGCCTCACC	CCTTTGTCTC	CAACCTCCAT	3600
TTCTGTTTAC	TTTGTGGAGA	GGGCATTACT	TGTTTCGTTAT	AGACATGGAC	GTTAAGAGAT	3660
ATTCAAAACT	CAGAAGCATC	AGCAATGTTT	CTCTTTTCTT	AGTTTATTCT	GCAGAATGGA	3720
AACCCATGCC	TATTAGAAAT	GACAGTACTT	ATTAATTGAG	TCCCTAAGGA	ATATTACAGC	3780
CACTACATAG	ATAGCTTTTT	TTTTTTTTTT	TTTTTTTTTA	TAAGGACACC	TCTTTCCAAA	3840
CAGGCCATCA	AATATGTTCT	TATCTCAGAC	TTACGTTGTT	TTAAAAGTTT	GGAAAGATAC	3900
ACATCTTTTC	ATACCCCCCC	TTAGGAGGTT	GGGCTTTTCT	ATCACCTCAG	CCAACCTGTG	3960
CTCTTAATTT	ATTGCATAAT	GATATCCACA	TCAGCCAAC	GTGGCTCTTT	AATTTATTGC	4020
ATAATGATAT	TCACATCCCC	TCAGTTGCAG	TGAATTGTGA	GCAAAAAGATC	TTGAAAGCAA	4080
AAAGCACTAA	TTAGTTTAAA	ATGTCACCTT	TTTGGTTTTT	ATTATACAAA	AACCATGAAG	4140
TACTTTTTTT	ATTTGCTAAA	TCAGATTGTT	CTTTTTTAGT	GACTCATGTT	TATGAAGAGA	4200
GTTGAGTTTA	ACAATCCTAG	CTTTTAAAAG	AAACTATTTA	ATGTAAAATA	TTCTACATGT	4260
CATTGAGATA	TTATGTATAT	CTTCTAGCCT	TTATCTGTGA	CTTTAATGT	ACATATTTCT	4320
GTCTTGCGTG	ATTTGTATAT	TTCAGTGGTT	TAAAAACAA	ACATCGAAAG	GCTTATTCCA	4380
AATGGAAGAT	AGAATATAAA	ATAAAACGTT	ACTTGTAATA	AAAAAAA		

Seq ID NO: 249 Protein sequence:  
Protein Accession #: NP\_003383

85  
1 11 21 31 41 51



MAGSAMSSKF FLVALAIFFS FAQVVEANS WWSLGMNPNV QMSEVYIIGA QPLCSQLAGL 60  
SQGQKKLCHL YQDHMQYIGE GAKTGIKECQ YQFRHRRWNC STVDNNTSVFG RVMQIGSRET 120  
AFTYAVSAAG VVNAMSRAER EGELSTCGCS RAARPKDLPR DWLWGGCGDN IDYGYRFAKE 180  
FVDARERERI HAKGSYESAR ILMNLHNEA GRRTVYNLAD VACKCHGVSG SCSLKTCLWLQ 240  
LADFRKVGDA LKEKYDSAAA MRLNSRGKLV QVNSRFNSPT TQDLVYIDPS PDYCVRNES 300  
GSLGTQGRLC NKTSEGMDCG ELMCCGRGYD QFKTVQTERC HCKFWCCYV KCKKCTEIVD 360  
QFVCK

Seq ID NO: 250 DNA sequence  
Nucleic Acid Accession #: NM\_014058  
Coding sequence: 56..1324

1 11 21 31 41 51  
| | | | |  
15 TGACTTGAT GTAGACCTCG ACCTTCACAG GACTCTTCAT TGCTGGTTGG CAATGATGTA 60  
TCGCCAGAT GTGGTGAGGG CTAGGAAAAG AGTTTGTGG GAACCTGGG TTATCGGCCT 120  
CGTCATCTTC ATATCCCTGA TTGTCTGGC AGTGTGCATT GGACTCACTG TTCATTATGT 180  
GAGATATAAT CAAAAGAAGA CCTACAATTA CTATAGCACA TTGTCAATTA CAACTGACAA 240  
ACTATATGCT GAGTTTGGCA GAGAGGCTTC TAACAATTTT ACAGAAATGA GCCAGAGACT 300  
20 TGAATCAATG GTGAAAATG CATTTTATAA ATCTCCATTA AGGGAAGAAT TTGTCAAGTC 360  
TCAGGTTATC AAGTTCAGTC AACAGAAGCA TGGAGTGTG GCTCATATGC TGTGATTG 420  
TAGATTTTAC TCTACTGAGG ATCCTGAAAC TGATAGATAA ATTGTTCAAC TTGTTTTACA 480  
TGAAAAGCTG CAAGATGCTG TAGGACCCCC TAAAGTAGAT CCTCACTCAG TTAATAATTAA 540  
25 AAAATCAAC AAGACAGAAA CAGACAGCTA TCTAAACCAT TGCTGCGGAA CACGAAGAAG 600  
TAAACTCTA GGTCAGAGTC TCAGGATCGT TGGTGGGACA GAAGTAGAAG AGGGTGAATG 660  
GCCCTGGCAG GCTAGCCTGC AGTGGGATGG GAGTCATCGC TGTGGAGCAA CCTTAATTAA 720  
TGCCACATGG CTTGTGAGTG CTGCTCACTG TTTTACAACA TATAAGAACC CTGCCAGATG 780  
GACTGCTTCC TTTGGAGTAA CAATAAAACC TTCGAAAATG AAACGGGGTC TCCGGAGAAT 840  
30 AATTGTCCAT GAAAAATACA AACACCCATC ACATGACTAT GATATTTCTC TTGCAGAGCT 900  
TTCTAGCCCT GTTCCCTACA CAAATGCAGT ACATAGAGTT TGTCTCCCTG ATGCATCCTA 960  
TGAGTTTCAA CCAGGTGATG TGATGTTTGT GACAGGATTT GGAGCACTGA AAAATGATGG 1020  
TTACAGTCAA AATCATCTTC GACAAGCACA GGTGACTCTC ATAGACGCTA CAACTTGCAA 1080  
TGAACCTCAA GCTTACAATG ACGCCATAAC TCCTAGAATG TTATGTGCTG GCTCCTTAGA 1140  
AGGAAAAACA GATGCATGCC GAGGTGACTC TGGAGGACCA CTGGTTAGTT CAGATGCTAG 1200  
35 AGATATCTGG TACCTTGCTG GAATAGTGAG CTGGGGAGAT GAATGTGCGA AACCACAAA 1260  
GCCTGGTGTT TATACTAGAG TTACGGCCTT GCGGGACTGG ATTACTTCAA AAAGTGGTAT 1320  
CTAAGAGAGA AAAGCCTCAT GGAACAGATA ACATTTTTTT TTGTTTTTTT GGTGTGGAGG 1380  
CCATTTTTAG AGATACAGAA TTGGAGAAGA CTTGCAAAAC AGCTAGATTT GACTGATCTC 1440  
40 AATAAACTGT TTGCTTGATG CAAAAA AAA A

Seq ID NO: 251 Protein sequence:  
Protein Accession #: NP\_054777

1 11 21 31 41 51  
| | | | |  
45 MYRPDVVRAR KRCWEPWVI GLVIFISLIV LAVCIGLTVH YVRYNQKITY NYYSTLSFTT 60  
DKLYAEFGRE ASNNFTEMSQ RLESMVKNFA YKSPLEEFV KSQVIKFSQ KHGVLAMHLL 120  
50 ICRPHSTEDP ETVDKIVQLV LHEKLQDAVG PPKVDPHSVK IKKINKTETD SYLNHCCGTR 180  
RSKTLGQSLR IVGGTEVEEG EWPWQASLQW DGSRRGATL INATWLVSAA HCFPTYKNPA 240  
RWTASFGVTI KPSKMKRGLR RIIVHEKYKH PSHDYDISLA ELSSPVPTYN AVHRVCLPDA 300  
SYEFQPGDVM FVTGFGALKN DGYSQNHRLQ AQVTLIDATT CNEPQAYNDA ITPRMLCAGS 360  
LEGKTDACQG DSGGPLVSSD ARDIWYLAGI VSWGDECAKP NKPGVYTRVT ALRDWITSKT 420  
GI

Seq ID NO: 252 DNA sequence  
Nucleic Acid Accession #: NM\_003504.2  
Coding sequence: 71-1771

1 11 21 31 41 51  
| | | | |  
60 GGCACGAGGC CTCGTGCCGC CGGGCTCTTG GTACCTCAGC GCGAGCGCCA GCGGTCCGGC 60  
CGCCGTGGCT ATGTTCTGTG CCGATTTCCG CAAAGAGTTC TACGAGGTGG TCCAGAGCCA 120  
GAGGGTCCCT CTCTTCGTGG CCTCGGACGT GGATGCTCTG TGTGCGTGCA AGATCCTTCA 180  
65 GGCTTGTTC CAGTGTGACC ACGTGCAATA TACGCTGGTT CCAGTTTCTG GGTGGCAAGA 240  
ACTTGAAACT GCATTTCTTG AGCATAAAGA ACAGTTTCAT TATTTTATTC TCATAAACTG 300  
TGGAGCTAAT GTAGACCTAT TGGATATTCT TCAACCTGAT GAAGACACTA TATCTTTTGT 360  
GTGTGACACC CATAGGCCAG TCAATGTCTG CAATGTATAC AACGATACCC AGATCAAATT 420  
70 ACTCATTAAA CAAGATGATG ACCTTGAAGT TCCCGCCTAT GAAGACATCT TCAGGGATGA 480  
AGAGGAGGAT GAAGAGCATT CAGGAAATGA CAGTGATGGG TCAGAGCCTT CTGAGAAGCG 540  
CACACGGTTA GAAGAGGAGA TAGTGGAGCA AACCATGCGG AGGAGGCAGC GGCGAGAGTG 600  
GGAGGCCCGG AGAAGAGACA TCCTCTTTGA CTACGAGCAG TATGAATATC ATGGGACATC 660  
GTCAGCCATG GTGATGTTTG AGCTGGCTTG GATGCTGTCC AAGGACCTGA ATGACATGCT 720  
75 GTGGTGGGCC ATCGTTGGAC TAACAGACCA GTGGGTGCAA GACAAGATCA CTCAAATGAA 780  
ATACGTGACT GATGTTGGTG TCCTGCAGCG CCACGTTTCC CGCCACAACC ACCGGAACGA 840  
GGATGAGGAG AACACACTCT CCGTGGACTG CACACGGATC TCCTTTGAGT ATGACCTCCG 900  
CCTGGTGCTC TACCAGCACT GGTCCCTCCA TGCAACACCA GCTATACCGC 960  
AGCCAGGTTT AAGCTGTGGT CTGTGCTGAG CACAGAAGCG CTCCAGGAGT TCCTTGACGA 1020  
80 CATGGGTCTT CCCCTGAAGC AGGTGAAGCA GAAGTTCAG GCCATGGACA TCTCCTTGAA 1080  
GGAGAAATTG CGGGAATGA TTGAAGAGTC TGCAAATAAA TTTGGGATGA AGGACATGCG 1140  
CGTGACAGCT TTCAGCATT ATTTTGGGTT CAAGCACAAG TTTCTGGCCA GCGACGTGGT 1200  
CTTTGCCACC ATGTCTTTGA TGGAGAGCCC CGAGAAGGAT GGCTCAGGGA CAGATCACTT 1260  
CATCCAGGCT CTGGACAGCC TCTCCAGGAC TAACCTGGAC AAGCTGTACC ATGGCCTGGA 1320  
85 ACTCGCCAAG AAGCAGCTGC GAGCCACCCA GCAGACCATT GCCAGCTGCC TTGACACCAA 1380  
CCTCGTCATC TCCAGGGGCT CTTTCTGTGA CTGCTCTCTC ATGGAGGGCA CTCCAGATGT 1440  
CATGCTGTTT TCTAGGCCGG CATCCCTAAG CCTGCTCAGC AAACACCTGC TCAAGTCCTT 1500  
TGTGTGTTTC ACAAAGAACC GCGCTGCAA ACTGCTGCCC CTGGTGATGG CTGCCCCCTT 1560



GAGCATGGAG CATGGCACAG TGACCGTGGT GGGCATCCCC CCAGAGACCG ACAGCTCGGA 1620  
 CAGGAAGAAC TTTTGTGGGA GGGCGTTTGA GAAGGCAGCG GAAAGCACCA GCTCCCGGAT 1680  
 GCTGCACAAC CATTGTGACC TCTCAGTAAT TGAGCTGAAA GCTGAGGATC GGAGCAAGTT 1740  
 TCTGGACGCA CTTATTTCCC TCCTGTCTTA GGAATTTGAT TCTTCCAGAA TGACCTTCTT 1800  
 ATTTATGTAA CTGGCTTTCA TTTAGATTGT AAGTTATGGA CATGATTTGA GATGTAGAAG 1860  
 CCATTTTFTA TTAAATAAAA TGCTTATTTT AGGCTCCGTC CCAAAAAAAA AAAAAAAA 1920  
 AAAAAAAA AA

Seq ID NO: 253 Protein sequence:  
 Protein Accession #: NP\_003495.1

1 11 21 31 41 51  
 | | | | |  
 MFVSDFRKEF YEYVQSQRVL LFAVDVDAL CACKILQALF QCDHVQYTLV PVSQWQLEET 60  
 AFLEHKEQFH YFILINGAN VDLLDILQPD EDTIFFVCDT HRPVNVVNVY NDTQIKLLIK 120  
 QDDLEVPAY EDIFRDEED EHSNGSDSG SEPSEKTRL EEEIVEQTMRRRQRREWEAR 180  
 RRDILFDYEQ YEYHGTSSAM VMFELAWMLS KDLNDMLWWA IVGLTDQWVQ DKITQMKYVT 240  
 DVGVLQRHVS RHNHRNEDEE NTLSVDCTRI SFEYDLRLVL YQHWSLHDSL CNTSYTAARF 300  
 KLWSVHGQKR LQEFLLADMGL PLKQVKQKFQ AMDISLKENL REMIEESANK FGMKDMRVQT 360  
 FSIHFGFKHK FLASDVVFAT MSLMESPEKD GSGTDHFIQA LDSLSRSNLD KLYHGLELAK 420  
 KQLRATQQT I ASCLCTNLVI SQGPFLYCSL MEGTPDVMLF SRPASLSLLS KHLKSFVCS 480  
 TKNRRCKLLP LVMAAPLSME HGTVTVVGIP PETDSSDRKN FFGRAFEKAA ESTSSRMLHN 540  
 HFDLSVIELK AEDRSKFLDA LISLLS

Seq ID NO: 254 DNA sequence  
 Nucleic Acid Accession #: NM\_022337  
 Coding sequence: 48..683

1 11 21 31 41 51  
 | | | | |  
 GGCTGCGCTT CCCTGGTCAG GCACGGCACG TCTGGCCGGC CGCCAGGATG CAGGCCCCGC 60  
 ACAAGGAGCA CCTGTACAAG TTGCTGGTGA TTGGCGACCT GGGCGTGGGG AAGACCAGTA 120  
 TCATCAAGCG CTACGTGCAC CAGAACTTCT CCTCGCACTA CCGGGCCACA ATCGGCGTGG 180  
 ACTTCGCGCT CAAGGTGCTC CACTGGGACC CGGAGACTGT GGTGCGCCTG CAGCTCTGGG 240  
 ATATCGCAGG TCAAGAAAGA TTTGGAAACA TGACGAGGGT CTATTACCGA GAAGCTATGG 300  
 GTGCATTTAT TGTCTTCGAT GTCACCAGGC CAGCCACATT TGAAGCAGTG GCAAAGTGGA 360  
 AAAATGATTT GGAATCCAAG TTAAGTCTCC CTAATGGCAA ACCGGTTTCA GTGGTTTTGT 420  
 TGGCCAACAA ATGTGACCAG GGAAGGATG TGCTCATGAA CAATGGCCTC AAGATGGACC 480  
 AGTTCTGCAA GGAGCACGGT TTCGTAGGAT GGTGTGAAAC ATCAGCAAAG GAAAATATAA 540  
 ACATTGATGA AGCCTCCAGA TGCCTGGTGA AACACATACT TGCAAATGAG TGTGACCTAA 600  
 TGGAGTCTAT TGAGCCGGAC GTCGTGAAGC CCCATCTCAC ATCAACCAAG GTTGCCAGCT 660  
 GCTCTGGCTG TGCCAAATCC TAGTAGGCAC CTTTGTCTGT GTCTGGTAGG AATGACCTCA 720  
 TTGTTCCACA AATTGTGCTT CTATTTTAC CATTTTGGGT AAACGTCAGG ATAGATATAC 780  
 CACATGTGGC AAGCCAAAGA TCTATGCCTC TGTTTTTCA ATGAGAGAGA AATAGCAAAT 840  
 GTTCTTTCTA TGCTTCTCTC ACCATCATCA CAGTCTTTAC AAACCTTTGA AAATATTTAG 900  
 TCTGTTACAA ACTTCTGTCA TGTAGTGAC CAAAATCCTG CAGGGCCACA GTCGGCACTG 960  
 TTATTTGCTT CTTTAAATCA GCAAAGGCCT CAAGTCTTAA AATAAAAGGG GAGAAGAACA 1020  
 AACTAGCTGT CAAGTCAAGG ACTGGCTTTC ACCTTGCCCT GGTGTCTTTT TCCAGATTTT 1080  
 AATATATTCT CTGATGGCCT GACAGGCCTA TTAAGTAGAT GTGATATTTT CTTCCAAGAT 1140  
 GACCTCCATT CTCGGCAGAC CTAAGAGTTG CCTCTGAGTT AGCTCTTTGG AATCGTGAAC 1200  
 ACAGGTGTGC TATATTGTCC TTGTCTTAAC TGTCACTTGC CATGGCCTGA ATGTTGGCTT 1260  
 AACTGAATAT TGTATGAAA GACATGCCTC CATATGTGCC TTTCTGTTAG CTCTCTTTGA 1320  
 CTCAAGCTGT GGGGCTCCTC TATACATGCT ATACATGTAA TATATATTAT ATATATTTTT 1380  
 GCAAGTGAAC AATAAAACAT TAAAAGATAA AA

Seq ID NO: 255 Protein sequence:  
 Protein Accession #: NP\_071732

1 11 21 31 41 51  
 | | | | |  
 MQAPHKEHLY KLLVIGDLGV GKTSIIKRYV HQNFSSHYRA TIGVDFALKV LHWDPEVTVR 60  
 LQLWDIAGQE RFGNMTRVYY REAMGAFIV DVTRPATFEA VAKWKNDLDS KLSLPNGKPV 120  
 SVVLLANKCD QGKDVLMNNG LKMDQFCKEH FVGWGFETSA KENINIDEAS RCLVKHILAN 180  
 ECDLMESIEP DVVKPHLTST KVASCSCGAK S

Seq ID NO: 256 DNA sequence  
 Nucleic Acid Accession #: NM\_016321  
 Coding sequence: 25..1464

1 11 21 31 41 51  
 | | | | |  
 GGAACCGCCC GCTGCCAGCC CGGCCAGGCA CCCCTGCAGC ATGGCCTGGA ACACCAACCT 60  
 CCGCTGGCGG CTGCCGCTCA CCTGCCTGCT CCTGCAGGTG ATTATGGTGA TTCTCTTCGG 120  
 GGTGTTTCGTG CGCTACGACT TCGAGGCCGA CGCCCACTGG TGGTCAGAGA GGACGCACAA 180  
 GAACCTGAGC GACATGGAGA ACGAATTCTA CTATCGCTAC CCAAGCTTCC AGGACGTGCA 240  
 CGTGATGGTC TTCGTGGGCT TCGGCTTCTT CATGACTTTC CTGCAGCGCT ACGGCTTCAG 300  
 CGCCGTGGGC TTCAACTTCC TGTGCGCAGC CTTGCGCATC CAGTGGGCGC TGCTCATGCA 360  
 GGGCTGGTTC CACTTCTTAC AAGACCGCTA CATCGTCTGT GGCGTGGAGA ACCTCATCAA 420  
 CGCTGACTTC TGCGTGGCCT CTGTCTGCGT GGCCTTTGGG GCAGTCTCTG GTAAAGTCAG 480  
 CCCCATTCAG CTGCTCATCA TGACTTTCTT CCAAGTGACC CTCTTCGCTG TGAATGAGTT 540  
 CATTCTCCTT AACCTGCTAA AGGTGAAGGA TGCAGGAGGC TCCATGACCA TCCACACATT 600  
 TGGCGCCTAC TTTGGGCTCA CAGTGACCCG GATCCTCTAC CGACGCAACC TAGAGCAGAG 660  
 CAAGGAGAGA CAGAA'TTCTG TGTACCAGTC GGACCTCTTT GCCATGATTG GCACCTCTT 720  
 CCTGTGGATG TACTGGCCCA GCTTCAACTC AGCCATATCC TACCATGGGG ACAGCCAGCA 780  
 CCGAGCCGCC ATCAACACCT ACTGCTCCTT GGCAGCCTGC GTGCTTACCT CCGTGGCAAT 840



ATCCAGTGCC CTGCACAAGA AGGGCAAGCT GGACATGGTG CACATCCAGA ATGCCACGCT 900  
CGCAGGAGGG GTGGCCGTGG GTACCGCTGC TGAGATGATG CTCATGCCTT ACGGTGCCCT 960  
CATCATCGGC TTCGTCTGCG GCATCATCTC CACCCTGGGT TTTGTATACC TGACCCCAT 1020  
CCTGGAGTCC CGGCTGCACA TCCAGGACAC ATGTGGCATT AACAATCTGC ATGGCATTCC 1080  
5 TGGCATCATA GGCGGCATCG TGGGTGCTGT GACAGCGGCC TCCGCCAGCC TTGAAGTCTA 1140  
TGGAAAAGAA GGGCTTGTCC ATTCCTTTGA CTTTCAAGGT TTCAACGGGG ACTGGACCGC 1200  
AAGAACACAG GGAAAGTTCC AGATTTATGG TCTCTTGGTG ACCCTGGCCA TGGCCCTGAT 1260  
GGGTGGCATC ATTTGTGGGC TCATTTTGAG ATTACCATTC TGGGGACAAC CTTCAGATGA 1320  
10 GAACTGCTTT GAGGATGCGG TCTACTGGGA GATGCCTGAA GGGAACAGCA CTGTCTACAT 1380  
CCCTGAGGAC CCCACCTTCA AGCCCTCAGG ACCCTCAGTA CCCTCAGTAC CCATGGTGTC 1440  
CCCACTACCC ATGGCTTCTT CGGTACCTT GGTACCTTAG GCTCCCAGGG CAGGTGAGGA 1500  
GCAGGCTCCA CAGACTSTCC TGGGGCCCCAG AGGAGCTGGT GCTGACCTAG CTAGGGATGC 1560  
AAGAGTGAGC AAGCAGCACC CCCACCTGCT GGCTTGGCCT CAAGGTGCCT CCACCCCTGC 1620  
CCTCCCCTTC ATCCCAGGGG GTCTGMCCTGA GAATGGAGAA GGAGAAGCTA CAAAGTGGGC 1680  
15 ATCCAAGCCG GGTCTTGGCT GCAGAAGTTC TGCTCTGCC TGGGGTCTTG GCCACATTGG 1740  
AGAAAAACAG GCTCAAAGTG GGGCTGGGAC CTGGTGGGTG AACCTGAGCT CTCCCAGGAG 1800  
ACAACTTAGC TGCCAGTCAC CACCTATGAG GCTCTTCTAC CCCGTGCCTG CACCTCGGCC 1860  
AGCATCTCCT ATGCTCCCTG GGTCCCCCAG ACCTCTCTGT GTTGTGTGCG TGGCAGCCTC 1920  
20 CAGGAATAAA CATTCTTGTT GTCCTTTGTA AAAAAAAAAA AAAAAAAA

Seq ID NO: 257 Protein sequence:  
Protein Accession #: NP\_057405

1 11 21 31 41 51  
| | | | |  
MAWNTNLRWR LPLTCLLLQV IMVILFGVVF RYDFEADAHW WSERTHKNLS DMENEFYRY 60  
PSFQDVHVMV FVGFGLMTF LQRYGFSAVG FNFLLAFFGI QWALLMQGWF HFLQDRYIV 120  
GVENLINADF CVASVCVAFG AVLKGVSPIQ LLIMTFFQVT LFAVNEFILL NLLKVKDAGG 180  
SMTIHTFGAY FGLTVTRILY RRNLEQSKER QNSVYQSDLF AMIGTLFLWM YWPSFNSAIS 240  
30 YHGDQSHRAA INTYCSLAAC VLTSAISSA LHKKGKLDLV HIQNLATLAGG VAVGTAAEMM 300  
LMPYGALIIG FVCGIISTLG FVYLTPFLES RLHIQDTCGI NNLHGIPGII GGIVGAVTAA 360  
SASLEVYKKE GLVHSFDFQG FNGDWTARTQ GKFIQYGLLV TLMALMGGI IVGLILRLPF 420  
WGQPSDENC F EDAVYWEMPE GNSTVIYIPED PTFKPSGPSV PSVPMVSLP MASSVPLVP

Seq ID NO: 258 DNA sequence  
Nucleic Acid Accession #: NM\_002358.2  
Coding sequence: 75..692

1 11 21 31 41 51  
| | | | |  
GGGAAGTGCT GTTGAGCCG CTGTGGTTGC TGTCCGCGGA GTGGAAGCGC GTGCTTTTGT 60  
TTGTGTCCCT GGCCATGGCG CTGCAGCTCT CCCGGGAGCA GGGAAATCACC CTGCGCGGGA 120  
GCGCCGAAAT CGTGGCCGAG TTCTTCTCAT TCGGCATCAA CAGCATTTTA TATCAGCGTG 180  
45 GCATATATCC ATCTGAAACC TTTACTCGAG TGCAGAAATA CGGACTCACC TTGCTTGTA 240  
CTACTGATCT TGAGCTCATA AAATACCTAA ATAATGTGGT GGAACAACCTG AAAGATTGGT 300  
TATACAAGTG TTCAGTTCAG AAATGGTTG TAGTTATCTC AAATATTGAA AGTGGTGAGG 360  
TCCTGGAAAG ATGGCAGTTT GATATTGAGT GTGACAAGAC TGCAAAAGAT GACAGTGCAC 420  
CCAGAGAAAA GTCTCAGAAA GCTATCCAGG ATGAAATCCG TTCAGTGATC AGACAGATCA 480  
50 CAGCTACGGT GACATTTCTG CCACTGTTGG AAGTTTCTTG TTCATTTGAT CTGCTGATTT 540  
ATACAGACAA AGATTTGGTT GTACCTGAAA AATGGGAAGA GTCGGGACCA CAGTTTATTA 600  
CCAATTCTGA GGAAGTCCGC CTTCGTTTAT TTAATACTAC AATCCACAAA GTAAATAGCA 660  
TGGTGGCCTA CAAAATTCCT GTCAATGACT GAGGATGACA TGAGGAAAAT AATGTAATTG 720  
TAATTTTGAA ATGTGGTTTT CCTGAAATCA GGTCATCTAT AGTTGATATG TTTTATTTCA 780  
55 TTGGTTAATT TTTACATGGA GAAAACCAA ATGATACTTA CTGAACTGTG TGTAATTGTT 840  
CCTTTATTTT TTTGGTACCT ATTTGACTTA CCATGGAGTT AACATCATGA ATTTATTGCA 900  
CATTTGTTCA AAGGAACCAG GAGGTTTTTT TGTCAACATT GTGATGTATA TTCCTTTGAA 960  
GATAGTAAC TTAGATGGAA AAATCTGTGC TATAAAGCTA GATGCTTTCC TAAATCAGAT 1020  
GTTTTGGTCA AGTAGTTTGA CTCAGTATAG GTAGGGAGAT ATTTAAGTAT AAAATACAAC 1080  
60 AAAGGAAGTC TAAATATTCA GAATCTTTGT TAAGGTCCTG AAAGTAACTC ATAATCTATA 1140  
AACAATGAAA TATTGCTGTA TAGCTCCTTT TGACCTTCAT TTCATGTATA GTTTTCCCTA 1200  
TTGAATCAGT TTCCAATTAT TTGACTTTAA TTTATGTAAC TTGAACCTAT GAAGCAATGG 1260  
ATATTTGTAC TGTTAATGT TCTGTGATAC AGAACTCTTA AAAATGTTTT TTCATGTGTT 1320  
65 TTATAAATC AAGTTTTAAG TGAAAGTGAG GAAATAAAGT TAAGTTTGTT TTAATAAAAA 1380  
AAAAAAAAA

Seq ID NO: 259 Protein sequence:  
Protein Accession #: NP\_002349.1

1 11 21 31 41 51  
| | | | |  
MALQLSREQG ITLRGSABIV AEFFSFGINS ILYQRGIYPS ETFTRVQKYG LTLVTTDLE 60  
75 LIKYLNNVVE QLKDWLYKCS VQKLVVVISN IESGEVLERW QFDIECDKTA KDDSAPREKS 120  
QKAIQDEIRS VIRQITATVT FLPLLEVSCS FDLIYTDKD LVVPEKWEES GPQFITNSEE 180  
VRLRSFTTTI HKVNSMVAYK IPVND

Seq ID NO: 260 DNA sequence  
Nucleic Acid Accession #: NM\_001211  
Coding sequence: 43..3195

1 11 21 31 41 51  
| | | | |  
AAAGGCCTGC AGCAGGACGA GGACCTGAGC CAGGAATGCA GGATGGCGGC GGTGAAGAAG 60  
85 GAAGGGGGTG CTCTGAGTGA AGCCATGTCC CTGGAGGGAG ATGAATGGGA ACTGAGTAAA 120  
GAAAATGTAC AACCTTTAAG GCAAGGGCGG ATCATGTCCA CGCTTCAGGG AGCACTGGCA 180  
CAAGAATCTG CCTGTAACAA TACTCTTCAG CAGCAGAAAC GGGCATTGTA ATATGAAATT 240



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

CGATTTTACA CTGGAAATGA CCCTCTGGAT GTTTGGGATA GGTATATCAG CTGGACAGAG 300  
CAGAACTATC CTCAAGGTGG GAAAGAGAGT AATATGTCAA CGTTATTAGA AAGAGCTGTA 360  
GAAGCACTAC AAGGAGAAAA ACGATATTAT AGTGATCCTC GATTTCTCAA TCTCTGGCTT 420  
AAATTAGGGC GTTTATGCAA TGAGCCTTTG GATATGTACA GTTACTTGCA CAACCAAGGG 480  
ATTGGTGTTC CACTTGCTCA GTTCTATATC TCATGGGCAG AAGAATATGA AGCTAGAGAA 540  
AACTTTAGGA AAGCAGATGC GATATTTTCAG GAAGGGATTG AACAGAAGGC TGAACCACTA 600  
GAAAGACTAC AGTCCCAGCA CCGACAATTC CAAGCTCGAG TGTCTCGGCA AACTCTGTTG 660  
GCACTTGAGA AAGAAGAAGA GGAGGAAGTT TTTGAGTCTT CTGTACCACA ACGAAGCACA 720  
CTAGCTGAAC TAAAGAGCAA AGGGAAGGAG ACAGCAAGAG CTCCAATCAT CCGTGTAGGA 780  
GGTGCTCTCA AGGCTCCAAG CCAGAACAGA GGACTCCAAA ATCCATTTCC TCAACAGATG 840  
CAAAATAATA GTAGAATTAC TGTTTTTTGAT GAAAATGCTG ATGAGGCTTC TACAGCAGAG 900  
TTGTCTAAGC CTACAGTCCA GCCATGGATA GCACCCCCCA TGCCCAGGGC CAAAGAGAAT 960  
GAGCTGCAAG CAGGCCCTTG GAACACAGGC AGGTCTTTGG AACACAGGCC TCGTGGCAAT 1020  
ACAGCTTCAC TGATAGCTGT ACCCGCTGTG CTTCCCAGTT TCACTCCATA TGTGGAAGAG 1080  
ACTGCACAAC AGCCAGTTAT AACCATGTTT AAAATTGAAC CTAGTATAAA CCACATCCTA 1140  
AGCACCAGAA AGCCTGGAAA GGAAGAAGGA GATCCTCTAC AAAGGGTTCA GAGCCATCAG 1200  
CAAGCGTCTG AGGAGAAGAA AGAGAAGATG ATGTATTGTA AGGAGAAGAT TTATGCAGGA 1260  
GTAGGGGAAT TCTCCTTTGA AGAAATTCGG GCTGAAGTTT TCCGGAAGAA ATTAAGAGAG 1320  
CAAAGGGAAG CCGAGCTATT GACCAGTGCA GAGAAGAGAG CAGAAATGCA GAAACAGATT 1380  
GAAGAGATGG AGAAGAAGCT AAAAGAAATC CAACTACTC AGCAAGAAAG AACAGGTGAT 1440  
CAGCAAGAAG AGACGATGCC TACAAAGGAG ACAAAATGAC TGCAAATTGC TTCCGAGTCT 1500  
CAGAAAATAC CAGGAATGAC TCTATCCAGT TCTGTTTGTG AAGTAACTG TTGTGCCAGA 1560  
GAAACTTCAC TTGCGGAGAA CATTTGGCAG GAACAACCTC ATTCTAAAGG TCCAGTGTA 1620  
CCTTTCTCCA TTTTGTATGA GTTCTTCTT TCAGAAAAGA AGAATAAAAG TCCTCCTGCA 1680  
GATCCCCCAC GAGTTTTAGC TCAACGAAGA CCCCTTGTCAG TTCTCAAAAC CTCAGAAAAGC 1740  
ATCACCTCAA ATGAAGATGT GTCTCCAGAT GTTTGTGATG AATTTACAGG AATTGAACCC 1800  
TTGAGCGAGG ATGCCATTAT CACAGGCTTC AGAAATGTAA CAATTTGTCC TAACCCAGAA 1860  
GACACTTGTG ACTTTGCCAG AGCAGCTCCT TTTGTATCCA CTCCTTTTCA TGAGATAATG 1920  
TCCTTGAAGG ATCTCCCTTC TGATCCTGAG AGACTGTTAC CGGAAGAAGA TCTAGATGTA 1980  
AAGACCTCTG AGGACCAGCA GACAGCTTGT GGCATATCT ACAGTCAGAC TCTCAGCATC 2040  
AAGAAGCTGA GCCCAATTAT TGAAGACAGT CGTGAAGCCA CACACTCCTC TGGCTTCTCT 2100  
GGTTCTTCTG CCTCGGTTGC AAGCACCTCC TCCATCAAAT GTCTTCAAAT TCCTGAGAAA 2160  
CTAGAACTTA CTAATGAGAC TTCAGAAAAC CCTACTCAGT CACCATGGTG TTCACAGTAT 2220  
CGCAGACAGC TACTGAAGTC CCTACGAGT TTAAGTGCTT CTGCAGAGTT GTGTATAGAA 2280  
GACAGACCAA TGCCTAAGTT GGAAATTGAG AAGGAAATTG AATTAGGTAA TGAGGATTAC 2340  
TGCAATTAAC GAGAATACCT AATATGTGAA GATTACAAGT TATTCTGGGT GGCGCCAAGA 2400  
AACTCTGCAG AATTAACAGT AATAAAGGTA TCTTCTCAAC CTGTCCCAGT GGACTTTTAT 2460  
ATCAACCTCA AGTTAAAGGA ACGTTTAAAT GAAGATTTTG ATCATTTTGT CAGCTGTTAT 2520  
CAATATCAAG ATGGCTGTAT TGTTTGGCAC CAATATATAA ACTGCTTCAC CCTCAGGAT 2580  
CTTCTCCAAC ACAGTGAATA TATTACCCAT GAAATAACAG TGTTGATTAT TTATAACCTT 2640  
TTGACAATAG TGGAGATGCT ACACAAAGCA GAAATAGTCC ATGGTGACTT GAGTCCAAGG 2700  
TGTCTGATTC TCAGAAACAG AATCCACGAT CCCTATGATT GTAACAAGAA CAATCAAGCT 2760  
TTGAAGATAG TGGACTTTTC CTACAGTGTG GACCTTAGGG TGCAGCTGGA TGTTTTACC 2820  
CTCAGCGGCT TTCGGACTGT ACAGATCCTG GAAGGACAAA AGATCCTGGC TAACTGTTCT 2880  
TCTCCCTACC AGGTAGACCT GTTTGGTATA GCAGATTTAG CACATTTACT ATTGTTCAAG 2940  
GAACACCTAC AGGTCTTCTG GGATGGGTCC TTCTGGAAAC TTAGCCAAA TATTCTGAG 3000  
CTAAAAGATG GTGAATTGTG GAATAAATTC TTTGTGCGGA TTCTGAATGC CAATGATGAG 3060  
GCCACAGTGT CTGTTCTTGG GGAGCTTGCA GCAGAAATGA ATGGGGTTTT TGACACTACA 3120  
TTCCAAAGTC ACCTGAACAA AGCCTTATGG AAGGTAGGGA AGTTAACTAG TCCTGGGGCT 3180  
TTGCTCTTTC AGTGAGCTAG GCAATCAAGT CTCACAGATT GCTGCCTCAG AGCAATGGTT 3240  
GTATTGTGGA ACACTGAAAC TGTATGTGCT GTAATTTAAT TTAGGACACA TTTAGATGCA 3300  
CTACCATTGC TGTCTACTT TTTGGTACAG GTATATTTTG ACGTCACTGA TATTTTAT 3360  
ACAGTGATAT ACTTACTCAT GGCCTTGTCT AACTTTTGTG AAGAATATT TTATTTCTAAA 3420  
CAGACTCATT ACAAATGGTT ACCTTGTAT TTAACCCATT TGTCTCTACT TTTCCCTGTA 3480  
CTTTTCCCAT TTGTAATTTG TAAAATGTT TCTTATGATC ACCATGTATT TTGTAAATAA 3540  
TAAAATAGTA TCTGTTAAAA AAAAAAAAAA AAAAAAAAAA AAA

Seq ID NO: 261 Protein sequence:  
Protein Accession #: NP\_001202

60  
65  
70  
75  
80

1 11 21 31 41 51  
| | | | |  
MAAVKKEGGA LSEAMSLEGD EWELSKENVQ PLRQGRIMST LQGALAQESA CNNTLQQQKR 60  
AFEYEIRFYT GNDPLDVWDR YISWTEQNYQ QGGKESNMST LLERAVEALQ GEKRYSDPR 120  
FLNLWLKLGR LCNEPLDMYS YLHNQIGVSV LAQFYISWAE EYEARENFRK ADAIFQEGIQ 180  
QKAEPLERLQ SQHRQFQARV SQTLLALEK EEEEEVFESS VPQRSTLAEK KSKGKKTARA 240  
PIIRVGGALK APSQNRGLQN PFPQQMQNNS RITVFDENAD EASTAELSKP TVQPWIAPPM 300  
PRAKENELQA GPWNTGRSLR HRPRGNTASL IAVPAVLPSF TPYVEETAQQ PVMTPECKIEP 360  
SINHILSTRK PGKEEGDPLQ RVQSHQQASE EKKEKMMYCK EKIYAGVGEF SFEEIRAEVF 420  
RKKLKEQREA ELLTSAEKRA EMQKQIEEME KKLKEIQTTQ QERTGDQEE TMPTKETTKL 480  
QIASESQKIP GMTLSSSVQC VNCCARETSL AENIWQEQPH SKGPSVPFSI FDEFLSEKK 540  
NKSPPADPPR VLAQRRPLAV LKTSSESITSN EDVSPDVCDE FTGIEPLSED AITGFRNVT 600  
ICPNPEDTCD FARAARFVST PFHEIMSLKD LPSDPERLLP EEDLDVKTSE DQQTACGTIY 660  
SQTLSIKKLS PIIEDSREAT HSSGFSGSSA SVASTSSIKC LQIPEKLELT NETSENPTQS 720  
PWCSQYRRQL LKSLPELSAS AELCIEDRPM PKLEIEKEIE LGNEDYCIKR EYLICEDYKL 780  
FWVAPRNSAE LTVIKVSSQP VPWDFYINLK LKERLNEDFD HFCSCYQYQD GCIVWHQYIN 840  
CFTLQDLLOH SEYITHEITV LIIYNLLTIF EMLHKABIVH GDLSRCLIL RNRIHDPYDC 900  
NKNNQALKIV DFSYSVDLRV QLDVFTLSGF RTVQILEGQK ILANCSSPYQ VDLFGIADLA 960  
HLLLFKEHLQ VFWDGSFWKL SQNISLKDGL ELWNKFFVRI LNANDEATVS VLGELAAEMN 1020  
GVFDTTFQSH LNKALWKVGK LTSPGALLFQ

Seq ID NO: 262 DNA sequence  
Nucleic Acid Accession #: NM\_003784  
Coding sequence: 365..1507

85  
1 11 21 31 41 51



	GTCTACTTAT	CAATAAGCAG	CTGCCTGTGC	AGAGTGCAGG	CTGCACCTTT	GGACAGCCTT 60
	TAAAACTGAA	TTCTCAGAAT	TTTAGAACAA	ATTTTGTCT	AGAAATGCTG	ACTTTGGTTC 120
5	ATTAGGTAGT	GGTAAACAG	GCTCCCTTCG	AAGCTCTCCT	TCATCACCTT	CCTAAGTGCA 180
	TGTACAGGGA	AGCTCTCCTT	CATCACCTTC	CTAAGTGCAT	GGGGGAAAAT	ACCTAGGGCT 240
	CAACAGTCTT	GAGAAGTGTG	GAAACATTTT	CTTTGTGAGT	GAGAACAGAT	CACCTAGAGA 300
	AAGGAAACCA	GATTCCCATC	ACTGCTTCTG	GGTATCAGAT	GCTAGCGCTG	CACTCCATTT 360
	TGCAATGGCC	TCCCTTGCTG	CAGCAAATGC	AGAGTTTTCG	TTCAACCTGT	TCAGAGAGAT 420
10	GGATGACAAAT	CAAGGAAATG	GAAATGTGTT	CTTTCTCTCT	CTGAGCCTCT	TCGCTGCCCT 480
	GGCCCTGGTC	CGCTTGGGCG	CTCAAGATGA	CTCCCTCTCT	CAGATTGATA	AGTTGCTTCA 540
	TGTTAACACT	GCCTCAGGAT	ATGGAAACTC	TTCTAATAGT	CAGTCAGGGC	TCCAGTCTCA 600
	ACTGAAAAGA	GTTTTTCTG	ATATAAATGC	ATCCCACAAG	GATTATGATC	TCAGCATTTG 660
	GAATGGGCTT	TTTGCTGAAA	AAGTGTATGG	CTTTCATAAG	GACTACATTG	AGTGTGCCGA 720
	AAAATTATAC	GATGCCAAAG	TGGAGCGAGT	TGACTTTACG	AATCATTTAG	AAGACACTAG 780
15	ACGTAATATT	AATAAGTGGG	TTGAAAATGA	AACACATGGC	AAAATCAAGA	ACGTGATTGG 840
	TGAAGGTGGC	ATAAGCTCAT	CTGCTGTAAT	GGTGTGGTG	AATGCTGTGT	ACTTCAAAGG 900
	CAAGTGGCAA	TCAGCCTTCA	CCAAGAGCGA	AACCATAAAT	TGCCATTTCA	AATCTCCCAA 960
	GTGCTCTGGG	AAGGCAGTCG	CCATGATGCA	TCAGGAACGG	AAGTTCAATT	TGCTGTAT 1020
20	TGAGGACCCA	TCAATGAAGA	TTCTTGAGCT	CAGATACAAT	GGTGGCATAA	ACATGTACGT 1080
	TCTGCTGCCT	GAGAATGACC	TCTCTGAAAT	TGAAAACAAA	CTGACCTTTC	AGAATCTAAT 1140
	GGAATGGACC	AATCCAAGGC	GAATGACCTC	TAAGTATGTT	GAGGTATTTT	TTCTCAGTT 1200
	CAAGATAGAG	AAGAATTATG	AAATGAAACA	ATATTTGAGA	GCCCTAGGGC	TGAAAGATAT 1260
	CTTTGATGAA	TCCAAAGCAG	ATCTCTCTGG	GATTGCTTCG	GGGGGTCGTC	TGTATATATC 1320
25	AAGGATGATG	CACAAATCTT	ACATAGAGGT	CACTGAGGAG	GGCACCAGAG	CTACTGCTGC 1380
	CACAGGAAGT	AATATTGTAG	AAAAGCAACT	CCCTCAGTCC	ACGCTGTTTA	GAGCTGACCA 1440
	CCCATTCCCTA	TTTGTATCA	GGAAGGATGA	CATCATCTTA	TTCACTGGCA	AAGTTTCTTG 1500
	CCCTTGAAAA	TCCAATTGGT	TTCTGTTATA	GCAGTCCCCA	CAACATCAAA	GRACCACCAC 1560
	AAGTCAATAG	AATYGRGTTT	AATTGGAAAA	ATGTGGTGT	TCCTTTGAGT	TTATTTCTTC 1620
30	CTAACATTGG	TCAGCAGATG	ACACTGGTGA	CTTGACCCTT	CCTAGACACC	TGGTTGATTG 1680
	TCCTGATCCC	TGCTCTTAGC	ATTCTACCAC	CATGTGTCTC	ACCCATTTCT	AATTTTATTG 1740
	TCTTTCTTCC	CACGCTCATT	TCTATCATTC	TCCCCCATGA	CCCGTCTGGA	AATTATGGAG 1800
	RGTGCTCAAC	TGGTAAGGAG	AACGTAGAAG	TAGCCCTAGG	GATCCTTTTT	GAAACTCTAC 1860
	AGTTATCGCA	GATATTCTAG	CTTCATTGTA	AGCAATCTAG	GAAATAAGCC	CTGCTGCTTT 1920
35	CTAGAAATAA	TGTGAAGGA	TAAATTTTCT	TTGTTGACCT	ATGAAGATTT	TAGAGTTTAC 1980
	CTTCATATGT	TTGATTTTAA	ATCAGTGTAT	AATCTAGATG	GTAAAAAATG	TGAAATTGGG 2040
	ATTAGGGACC	TACCAAAATA	TTTCATTAAT	GCTTTCAATT	GACAAATTTT	GGCCTTTCTT 2100
	TGATAAGACA	ATATGTACAT	GTTTTTTCAA	ATATTAAAGA	TCTTTTAACT	GTTGGCAGTT 2160
	GTTATCTACA	GAATCATATT	TCATATGCTG	TGTAGTTTAT	AAGTTTTTCC	TCTATTTATC 2220
40	AGAATAAAGA	AATACAACAT	ACCTGTAAA			

Seq ID NO: 263 Protein sequence:  
Protein Accession #: NP\_003775

45	1	11	21	31	41	51	
	MASLAAANAE	FCFNLFREMD	DNQNGNVPFF	SSLSLFAALA	LVRLGAQDD	LSQIDKLLHV	60
	NTASGYGNSS	NSQSLQSQL	KRVFSDINAS	HKDYDLSIVN	GLFAEKVYGF	HKDYIECAEK	120
50	LYDAKVERVD	FTNHLEDTRR	NINKWVENET	HGKIKNVIGE	GGISSAVMV	LVNAVYFKGK	180
	WQSAFTKSET	INCHFKSPKC	SGKAVAMMHQ	ERKFNLVIE	DPSMKILELR	YNGGINMYVL	240
	LPENDLSEIE	NKLTQNLME	WTNPRRMTSK	YVEVFFPQFK	IEKNYEMKQY	LRLGLKIDIF	300
	DESKADLSGI	ASGGRLYISR	MMHKSYYIEVT	EEGTEATAAT	GSNIVEKQLP	QSTLFRADHP	360
	FLFVIRKDDI	ILFSGKVSCP					

55 Seq ID NO: 264 DNA sequence  
Nucleic Acid Accession #: AB052906  
Coding sequence: 74-814

60	1	11	21	31	41	51	
	AAAACCTTGA	GGTGATTTCAT	CTTCCAGGCT	CTCCTTCCAT	CAAGTCTCTC	CTCCCTAGCG	60
	CTCTGGGTCC	TTAATGGCAG	CAGCCGCCGC	TACCAAGATC	CTTCTGTGCC	TCCCGCTTCT	120
	GCTCCTGCTG	TCCGGCTGGT	CCCGGGCTGG	GCGAGCCGAC	CCTCACTCTC	TTTGCTATGA	180
	CATCACCGTC	ATCCCTAAGT	TCAGACCTGG	ACCACGGTGG	TGTGCGGTTC	AAGGCCAGGT	240
65	GGATGAAAAG	ACTTTTCTTC	ACTATGATCG	TGGCAACAAG	ACAGTCACAC	CTGTCACTCC	300
	CCTGGGGAAG	AACTAAATG	TCACAACGGC	CTGGAAAGCA	CAGAACCCAG	TACTGAGAGA	360
	GGTGGTGGAC	ATACTTACAG	AGCAACTGCG	TGACATTTCAG	CTGGAGAATT	ACACACCCAA	420
	GGAACCCCTC	ACCCTGCAGG	CCAGGATGTC	TTGTGAGCAG	AAAGCTGAAG	GACACAGCAG	480
70	TGGATCTTGG	CAGTTCACTT	TCGATGGGCA	GATCTTCCTC	CTCTTTGACT	CAGAGAAGAG	540
	AATGTGGACA	ACGGTTCATC	CTGGAGCCAG	AAAGATGAAA	GAAAAGTGGG	AGAATGACAA	600
	GGTTGTGGCC	ATGTCCTTCC	ATTACTTCTC	AATGGGAGAC	TGTATAGGAT	GGCTTGAGGA	660
	CTTCTTGATG	GGCATGGACA	GCACCTGGA	GCCAAGTGCA	GGAGCACCAC	TCGCCATGTC	720
	CTCAGGCACA	ACCCAACCTCA	GGGCCACAGC	CACCACCCTC	ATCCTTTGCT	GCCTCCTCAT	780
75	CATCCTCCCC	TGCTTCATCC	TCCCTGGCAT	CTGAGGAGAG	TCCTTTAGAG	TGACAGGTTA	840
	AAGCTGATAC	CAAAAGGCTC	CTGTGAGCAC	GGTCTTGATC	AAACTCGCCC	TTCTGTCTGG	900
	CCAGCTGCCC	ACGACCTACG	GTGTATGTCC	AGTGGCCTCC	AGCAGATCAT	GATGACATCA	960
	TGGACCCCAAT	AGCTCATTCA	CTGCCTTGAT	TCCTTTTGCC	AACAATTTTA	CCAGCAGTTA	1020
	TACCTAACAT	ATTATGCAAT	TTTCTCTTGG	TGCTACCTGA	TGGAATTCCT	GCACTTAAAG	1080
	TTCTGGCTGA	CTAAACAAGA	TATATCATTT	TCTTTCTTCT	CTTTTGTGTT	GGAAAATCAA	1140
80	GTAATTCTTT	GAATGATGAT	CTCTTTCTTG	CAAATGATAT	TGTCAGTAAA	ATAATCACGT	1200
	TAGACTTCAG	ACCTCTGGGG	ATTCTTTCCG	TGTCCTGAAA	GAGAAATTTT	AAATTATTTA	1260
	ATAAGAAAAA	ATTTATATTA	ATGATTGTTT	CCTTTAGTAA	TTTATTTGTT	TGTACTGATA	1320
	TTTAAATAAA	GAGTTCTATT	TCCCAAAAAA	AAAAAAAAAA	A		

85 Seq ID NO: 265 Protein sequence:  
Protein Accession #: BAB61048.1



```

1      11      21      31      41      51
|      |      |      |      |      |
MAAAAATKIL LCLPLLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGQVDEKT 60
FLHYDCGNKT VTPVSPLGKK LNVTTAWKAQ NPVLREVVDI LTEQLRDIQL ENYTPKEPLT 120
5  LQARMSCEQK AEGHSSGSWQ FSGDGQIFLL FDSEKRMWTT VHPGARKMKE KWENDKVVAM 180
SFHYFSMGDC IGWLEDFLMG MDSTLEPSAG APLAMSSGTT QLRATATTLI LCCLLIILPC 240
FILPGI

```

Seq ID NO: 266 DNA sequence  
Nucleic Acid Accession #: XM\_084853.1  
Coding sequence: 127-444

```

1      11      21      31      41      51
|      |      |      |      |      |
ATTGATGATA TATTTAACGA AATCAAATTT GGTGAATATG TGGACACTGG AAAGCTAATC 60
GACAAGATCA ACTTACCAGA TTTCTTAAAA GTGTACCTTA ACCACAAGCC ACCTTTTGGT 120
AACACCATGA GTGGCATCCA CAAGAGCTTT GAGGTGCTCG GTTATACCAA CTCCAAAGGG 180
AAAAAGGCCA TTCGAAGAGA GGAAGTCTCTG AGACTGCTCG TTAATAAAGG TGAGCATATG 240
ACGGAGGAGG AGATGTTGGA TTGCTTTGCT TCACTGTTTG GCCTGAATCC CGAGGGATGG 300
20 AAATCCGAGC CTGCAACCTG CTCCGTCAAA GGTTCAGAAA TTTGCCTTGA AGAAGAACTT 360
CCAGACGAAA TCACTGCAGA AATATTGCGG ACTGAAATTC TTGGCTTAAC CATTTCAGAA 420
GATTCCCGCC AGGATGGTCA GTGAAGTTAC CAGGAATGTT TAAAGCACAA AGGACTTTGG 480
GTGTGTGTGC ATGCACATGT GTGTGTTTTT CATGAGGCAC TGCTTTTTTAT GCATTTCCCT 540
25 CCCCCCTCTC ATCTTTAGAA CATTTAGACA TTAAAGCAAG TTTCTGGTGA GCAATG

```

Seq ID NO: 267 Protein sequence:  
Protein Accession #: XP\_084853.1

```

1      11      21      31      41      51
|      |      |      |      |      |
MSGIHKSFV LGYTNSKGKK AIRREDFLRL LVTKEGHEMT EEMLD CFASL FGLNPEGWKS 60
EPATCSVKGS EICLEELPD EITAEIFATE ILGLTISED GQDQ

```

Seq ID NO: 268 DNA sequence  
Nucleic Acid Accession #: NM\_001898  
Coding sequence: 57-482

```

1      11      21      31      41      51
|      |      |      |      |      |
GGCTCTCACC CTCCTCTCCT GCAGCTCCAG CTTTGTGCTC TGCTCTGAG GAGACCATGG 60
CCCAGTATCT GAGTACCCTG CTGCTCCTGC TGGCCACCCT AGCTGTGGCC CTGGCCTGGA 120
GCCCCAAGGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCAATGATG 180
45 AGTGGGTACA GCGTGCCCTT CACTTCGCCA TCAGCGAGTA TAACAAGGCC ACCAAAGATG 240
ACTACTACAG ACGTCCGCTG CGGGTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGGTGA 300
ATTACTTCTT CGACGTAGAG GTGGGCCGCA CCATATGTAC CAAGTCCCAG CCCAACTTGG 360
ACACCTGTGC CTTCCATGAA CAGCCAGAAC TGCAGAAGAA ACAGTTGTGC TCTTTCGAGA 420
TCTACGAAGT TCCCTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCCT 480
AGGGATCTGT GCCAGGCCAT TCGCACCAGC CACCACCAC TCCCACCCC TGTAAGTGCTC 540
50 CCACCCCTGG ACTGGTGGCC CCCACCCTGC GGGAGGCCCT CCCATGTGCC TGCGCCAAGA 600
GACAGACAGA GAAGGCTGCA GGAGTCCTTT GTTGCTCAGC AGGGCGCTCT GCCCTCCCTC 660
CTTCCTTCTT GCTTCTAATA GCCCTGGTAC ATGGTACACA CCCCCCACC TCCTGCAATT 720
AAACAGTAGC ATCGCC

```

Seq ID NO: 269 Protein sequence:  
Protein Accession #: NP\_001889.1

```

1      11      21      31      41      51
|      |      |      |      |      |
MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWVQRALHF AISEYNKATK 60
DDYRRLPLRV LRARQQTGG VNYFFDVEVG RTICTKSQPN LDTCAFHEQP ELQKKQLCSF 120
EIYEVWENR RSLVKSRCQE S

```

Seq ID NO: 270 DNA sequence  
Nucleic Acid Accession #: XM\_093210  
Coding sequence: 13-1854

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGGCAAGCG CCGGAATCTC CTCAGCTGCC GTTTCACAAA AGAGGTACCA GGTCCGCACC 60
AAACGAGCAC ACAAGCAGCA CCAGAGCTG CAGAAGAAGG AGGCGGCAGC GATGGACCAG 120
GGCAGAGGGA ATGGGGAGGG GGCATCCTAC CCCATATCTG AGGTGCGACT GCGGGACGTA 180
GAGCGGACTG GGCCTTTCCC GTTGGCGCGT GGCCTCAATC AGGACTTCTT GCCCACGTGC 240
75 GCCTTCAAAA CGGTAAGAGC TGCAACTGAA CGTGTGAGAC ATGGTGAGCA TAGGCTGAGA 300
GGCGGCGGGA GAGATGCCCC TGAAGTCAAG TACCCGGACA CGCCCTCCAC TTCTACCACC 360
ACGAGTAACA CCGCCCCCAC GGGACCGCTC TCGAGGTCCC CCAAGCCAAG GACGCAAGGA 420
GGAACGCCCC GGCGCGCGGC CAGCAGCGGC GGGCACCAGG CCAATGGCCA CGGAAGTCAAG 480
CACTGGCAGT CGGCCCTCCT CACACCGCAG GCGTGCAAGT TGGCCGACGG AGCCTCCCGG 540
GCCGAGGACC CAGCTAGGCC GTCAACCCGG TTGCTCCAC GGAAGGGGGC ACCAGGCAAA 600
80 CTGCCCAAGG CCCCAGGCC AGGCTCCCTG GCGGAGGCCT CCGCTGGTCC CGCCAGATC 660
ATGGCCGCCA CCAAGCTCCC GAGCCATGGC TTCTGTCCG GGAACGGCCC GGCGTCTCTG 720
CTGTCCAGCT AG

```

Seq ID NO: 271 Protein sequence:  
Protein Accession #: XP\_093210

```

1      11      21      31      41      51

```



MLRHGEQKRR	RARKKWDFLP	TCAFKTVRAA	TERVRHGADR	LRGGGRDAHE	LKYPDTPSTS	60
TTTSNTAPTG	PLSRSPKPRT	QGGTPRRRPA	AAGTRANGHG	TQHWQSALLT	PQACSVADGA	120
SRAEDPARPS	PRLLPREGAP	GKLPKAPSPG	SLAEASAGLL	AHVRLQNADA	QRVSIQALP	180
PNSSVGRKEE	RPGAGQQRRA	PAPMATELST	GSRPSSHRRR	AVWPTEPPGP	RTQLEPSPRL	240
LPREGAPGKL	PKAPSPGSLA	EASAGPAQIM	AATRLPSRGF	LSGNPGASWL	SS	

Seq ID NO: 272 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..732

1	11	21	31	41	51	
GGATACTGTG	TCACTCAAAG	TAATGGGAGG	GAGAGAGAAC	AGGGAGGGTA	GGGATGCTTT	60
TGAAAAAGCT	TTTTTCCCA	CTTTTAACTT	GCTTTAGCGT	TAAGAGTACT	TACCAGCTAA	120
TAATGTGGAG	GAAATTATTC	TTTCTCATTG	GAGATTACAG	AATATATCTA	TTCATCTTGA	180
ATACCCACTT	GAAGCCTCTG	TAGAAATGTC	TCGTCCCTCCG	GTGTATTTTC	TAAAACCTAC	240
ATGATTTTGT	CTTGTTTCTG	CAGTGAGAAA	TTACATCCAT	AGCAAAGACA	AAAGTCTTTT	300
TAAATTATTT	TTATTTATCT	TTCATATAGT	TCTTACAATT	TCTAAAAAAT	TAACACTCAT	360
TTAGTATCAC	AATTTATGGG	AGAGGGTTTT	TTGTATTTTT	AAGCATATGT	GGCTTATATA	420
AAAATTGCAG	AAGTCATAGG	ACTGTCATGT	ATTGCAGCTC	TGAGAACCAA	TGCCTGAAAC	480
TTAAGCC						

Seq ID NO: 273 Protein sequence:  
Protein Accession #: Eos sequence

1	11	21	31	41	51	
MGGRENREGR	DAFEKAFFPT	FNLL				

Seq ID NO: 274 DNA sequence  
Nucleic Acid Accession #: NM\_003976.2  
Coding sequence: 299-961

1	11	21	31	41	51	
CTCTGAGCTT	CTCTGAGCCT	TGTTTGTCTCA	TCTGGAAAAA	GGGGATTAAA	CCATTTACCT	60
CATGGAGTTG	TGAAAGAATA	GCTGCAAAGC	ACCTAACACA	TAGTAAGGTT	CCCAGTGCAG	120
CTACTTCTGC	TGGGTTGAGT	CTAGCTGTGT	AGGCCCTTG	TTCTCACCT	GGAGAACTG	180
GGGTGGCAGG	CCGGTCCCC	ACAAAAGATA	ACTCATCTCT	TAATTTGCAA	GCTGCCTCAA	240
CAGGAGGGTG	GGGGAACAGC	TCAACAATGG	CTGATGGGCG	CTCCTGGTGT	TGATAGAGAT	300
GGAAC TTGGA	CTTGAGGACC	TCTCCACGCT	GTCCCACTGC	CCCTGGCCTA	GGCGGCAGCC	360
TGCCCTGTGG	CCCACCTGG	CCGCTCTGGC	TCTGCTGAGC	AGCGTCGCAG	AGGCCTCCCT	420
GGGCTCCGCG	CCCCGCAGCC	CTGCCCCCG	GGAAGGCCCC	CCGCCTGTCC	TGGCGTCCCC	480
CGCCCGCCAC	CTGCCGGGG	GACGCACGGC	CCGCTGGTGC	AGTGGAAAGAG	CCCGCGGGCC	540
GCCGCCGCAG	CCTTCTCGGC	CCGCGCCCC	GCCGCCTGCA	CCCCCATCTG	CTCTTCCCCG	600
CGGGGGCCGC	GCGGCGCGGG	CTGGGGGCCC	GGGCAGCCGC	GCTCGGGCAG	CGGGGGCGCG	660
GGGCTGCCGC	CTGCGCTCGC	AGCTGGTGCC	GGTGCGCGCG	CTCGGCCTGG	GCCACCGCTC	720
CGACGAGCTG	GTGCGTTTCC	GCTTCTGCGC	CGGCTCCTGC	CGCCGCGCGC	GCTCTCCACA	780
CGACCTCAGC	CTGGCCAGCC	TACTGGGCGC	CGGGGCCCTG	CGACCGCCCC	CGGGCTCCCG	840
GCCCGTCAGC	CAGCCCTGCT	GCCGACCCAC	GCGCTACGAA	GCGGTCTCCT	TCATGGACGT	900
CAACAGCACC	TGGAGAACCG	TGGACCGCCT	CTCCGCCACC	GCCTGCGGCT	GCCTGGGCTG	960
AGGGCTCGCT	CCAGGGCTTT	GCAGACTGGA	CCCTTACCGG	TGGCTCTTCC	TGCCTGGGAC	1020
CCTCCCGCAG	AGTCCCACTA	GCCAGCGGCC	TCAGCCAGGG	ACGAAGGCCT	CAAAGCTGAG	1080
AGGCCCTTAC	CGGTGGGTGA	TGGATATCAT	CCCCGAACAG	GTGAAGGGAC	AACTGACTAG	1140
CAGCCCCAGA	GCCCTCACCC	TGCGGATCCC	AGCCTAAAAG	ACACCAGAGA	CCTCAGCTAT	1200
GGAGCCCTTC	GGACCCACTT	CTCACAGACT	CTGGCACTGG	CCAGGCCTCG	AACCTGGGAC	1260
CCCTCCTCTG	ATGAACACTA	CAGTGGCTGA	GGCATCAGCC	CCCGCCAGG	CCCTGTAGGG	1320
ACAGCATTTG	AAGGACACAT	ATTGCAGTTG	CTTGGTTGAA	AGTGCCTGTG	CTGGAACCTG	1380
CCTGTACTCA	CTCATGGGAG	CTGGCCCC				

Seq ID NO: 275 Protein sequence:  
Protein Accession #: NP\_003967.1

1	11	21	31	41	51	
MELGLGGLST	LSHCPWPRRQ	PALWPTLAAL	ALLSSVAEAS	LGSAPRSPAP	REGPPPVLAS	60
PAGHLPGGRT	ARWCSGRARR	PPPQPSRPAP	PPPAPPSALP	RGGRAARAGG	PGSRARAAGA	120
RGCRRLRSQLV	PVRALGLGHR	SDELVRFRFC	SGSRRRARS	HDLSLASLLG	AGALRPPPGS	180
RPVSQPCCRP	TRYEAVSFMD	VNSTWRTVDR	LSATACGCLG			

Seq ID NO: 276 DNA sequence  
Nucleic Acid Accession #: NM\_057091.1  
Coding sequence: 783-1445

1	11	21	31	41	51	
ACTGGCCGCT	GAGAGAAGAA	TCGGGTGGAG	CAGAGAGCAG	CTGCTGCAGG	GCAGACAGCC	60
GGACCCCCAA	ATCTGCACGT	ACCAGCAGTC	AGCCGCCCCA	CGCAGGGACC	GGCTTACCCC	120
TCGCTCCCCG	CCCTCACTCA	CTTTCTCCCG	CCCTCGGCCC	GGCCTCCAG	CTCTCTACTT	180
CGCGTGCTTA	CAAACTCAAC	TCCCGGTTTC	CGTGCTCTC	CACCGCTCGA	GTTCTCTACT	240
CTCCATATCC	GAGGGGCCCC	TCCCGGTTTC	TACCCCTCTC	CCAACCTCGG	GGGACCTAGC	300
CAAGCTAGGG	GGGACTGGAT	CCGACGGGTG	GAGCAGCCAG	GTGAGCCCCG	AAAGGTGGGG	360
CGGGGCAGGG	GCGCTCCAG	CCCCACCCCG	GGATCTGGTG	ACGCTGGGGC	TGGAATTTGA	420
CACCGGACGG	CTGCGGCGGC	GGGCAGGAGG	CTGCTGAGGG	ATGGAGTTGG	GCCCGGCCCC	480
CAGACAAGGC	CCGGGGGCTC	CGCCAGCAGC	AGGTCCCTCG	GGCCCCAGCC	CTCGCTGCCA	540



5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60

```

CCCCGGCCTG GAGCCCCACA CCGGAGGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600
TAAAAGAGGC ACTGCCAGGT GTACAGTCCT GGGCATGCGC TGTTTGAGCT TCGGGGGAGA 660
GCCCAGCACT GGTCCCCGGA AAGGTGCCTA GAAGAACAAG GTGCAGGACC CCGTGCTGCC 720
TCAACAGGAG GGTGGGGGAA CAGCTCAACA ATGGCTGATG GCGCTCCTG GTGTTGATAG 780
AGATGGAAGT TGGACTTGGA GGCCTCTCCA CGCTGTCCCA CTGCCCCTGG CCTAGGCGGC 840
AGCCTGCCCT GTGGCCCCACC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCCT 900
CCCTGGGCTC CGCGCCCCGC AGCCTTGCCC CCGCGAAGG CCCCCGCTC GTCCTGGCGT 960
CCCCCGCCGG CCACCTGCCG GGGGGACGCA CGGCCCGCTG GTGCAGTGGA AGAGCCCGGC 1020
GGCCGCGGCC GCAGCCTTCT CGGCCGCGC CCGCGCCGCC TGACCCCCA TCTGCTCTTC 1080
CCCGCGGGGG CCGCGCGCGC CGGGCTGGGG GCGCGGGCAG CCGCGCTCGG GCAGCGGGGG 1140
CGCGGGGCTG CCGCTGCGC TCGCAGCTGG TGCCGGTGGC CGCGCTCGGC CTGGGCCACC 1200
GCTCCGACGA GCTGGTGCCT TTCCGCTTCT GCAGCGGCTC CTGCGCCCGC GCGCGCTCTC 1260
CACACGACCT CAGCCTGGCC AGCCTACTGG GCGCCGGGGC CCTGCGACCG CCCCCGGGCT 1320
CCCGGCCCCG CAGCCAGCCC TGCTGCCGAC CCACGCGCTA CGAAGCGGTC TCCTTCATGG 1380
ACGTCAACAG CACCTGGAGA ACCGTGGACC GCCTCTCCGC CACCGCCTGC GGCTGCCTGG 1440
GCTGAGGGCT CGCTCCAGGG CTTTGCAGAC TGGACCTTA CCGGTGGCTC TTCCTGCCTG 1500
GGACCTCCCC GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGACGAAG GCCTCAAAGC 1560
TGAGAGGCCC CTACCGGTGG GTGATGGATA TCATCCCCGA ACAGGTGAAG GGACAACCTGA 1620
CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG 1680
CTATGGAGCC CTTGCGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG 1740
GGACCCCTCC TCTGATGAAC ACTACATGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT 1800
AGGGACAGCA TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAAGTGCC TGTGCTGGAA 1860
CTGGCCTGTA CTCATCATG GGAGCTGGCC CC
  
```

Seq ID NO: 277 Protein sequence:

Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
 | | | | |  
 MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
 PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120  
 RGCRLRSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180  
 RPVSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 278 DNA sequence

Nucleic Acid Accession #: NM\_057160.1

Coding sequence: 1-714

40  
 45  
 50  
 55  
 60

```

1 11 21 31 41 51  

| | | | |  

ATGCCCCGCC TGATCTCAGC CCGAGGACAG CCCCTCCTTG AGGTCTCTCC TCCCCAAGCC 60  

CACCTGGGTG CCCTCTTTCT CCCTGAGGCT CCACTTGGTC TCTCCGCGCA GCCTGCCCTG 120  

TGGCCACCC TGGCCGCTCT GGCTCTGCTG AGCAGCGTCG CAGAGGCCTC CCTGGGCTCC 180  

GCGCCCCGCA GCCCTGCCCC CCGCGAAGGC CCGCCGCTG TCCTGGCGTC CCGCCCGGC 240  

CACCTGCCGG GGGGACGAC GCGCCGCTGG TGCAGTGGA GAGCCCGCG GCGCCGCGC 300  

CAGCCTTCTC GCGCCGCGCC CCGCCGCTG GCACCCCAT CTGCTCTTC CCGCGGGGC 360  

CGCGCGGCGC GGGCTGGGG CCGGGCAGC CGCGCTCGGG CAGCGGGGC GCGGGGCTGC 420  

CGCCTGCGCT CGCAGCTGGT GCGGTGCGC GCGCTCGGC TGGGCCACCG CTCCGACGAG 480  

CTGGTGCCTT TCCGCTTCTG CAGCGGCTCC TGCCGCGCG CGCGCTCTCC ACACGACCTC 540  

AGCCTGGCCA GCCTACTGGG CGCCGGGGCC CTGCGACCG CCGCGGCTC CCGGCCGTC 600  

AGCCAGCCCT GCTGCCGACC CACGCGCTAC GAAGCGGTCT CCTTCATGGA CGTCAACAGC 660  

ACCTGGAGAA CCGTGGACCG CCTCTCCGCC ACCGCTGCG GCTGCCTGG CTGAGGGCTC 720  

GCTCCAGGGC TTTGCAGACT GGACCTTAC CGGTGGCTCT TCCTGCCTGG GACCTCCCG 780  

CAGAGTCCCA CTAGCCAGCG GCCTCAGCCA GGGACGAAG CCTCAAAGCT GAGAGGCCCC 840  

TACCGGTGGG TGATGGATAT CATCCCCGAA CAGGTGAAG GACAACTGAC TAGCAGCCCC 900  

AGAGCCCTCA CCCTGCGGAT CCCAGCCTAA AAGACACCAG AGACCTCAG TATGGAGCCC 960  

TTCCGACCCA CTTCTCACAG ACTCTGGCAC TGGCCAGGCC TCGAACCTGG GACCCCTCCT 1020  

CTGATGAACA CTACAGTGGC TGAGGCATCA GCGCCCGCC AGGCCCTGTA GGGACAGCAT 1080  

TTGAAGGACA CATATTGCAG TTGCTTGGTT GAAAGTGCC GTGCTGGAAC TGGCCTGTAC 1140  

TCACTCATGG GAGCTGGCCC C
  
```

Seq ID NO: 279 Protein sequence:

Protein Accession #: NP\_476501.1

65  
 70

```

1 11 21 31 41 51  

| | | | |  

MPGLISARGQ PLLEVLPPQA HLGALFLPEA PLGLSAQPAL WPTLAALALL SVAEASLGS 60  

APRSPAPREG PPPVLASPAG HLPGGRTARW CSGRRARRPP QPSRPAPPPP APPSALPRGG 120  

RAARAGGPGS RARAAGARGC RLRSQLVFVR ALGLGHRSD LVRFRFCSGS CRRARSPHDL 180  

SLASLLGAGA LRPPGSRPV SQPCCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG
  
```

Seq ID NO: 280 DNA sequence

Nucleic Acid Accession #: NM\_057090.1

Coding sequence: 29-715

75  
 80  
 85

```

1 11 21 31 41 51  

| | | | |  

CTGATGGGCG CTCCTGGTGT TGATAGAGAT GGAACCTGGA CTTGGAGGCC TCTCCACGCT 60  

GTCCCACTGC CCCTGGCCTA GGCGGCAGGC TCCACTTGGT CTCTCCGCGC AGCCTGCCCT 120  

GTGGCCACCC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCTC CCCTGGGCTC 180  

CGCGCCCCGC AGCCCTGCCC CCGCGAAGG CCCCCGCTC GTCTTGGCT CCGCCCGGC 240  

CCACCTGCCG GGGGGACGCA CGGCCGCTG GTGCAGTGGA AGAGCCCGGC GCGCCCGCC 300  

GCAGCCTTCT CGGCCGCGC CCGCCGCGC TGCACCCCCA TCTGCTCTTC CCGCGGGGG 360  

CCGCGCGGCG CGGCTGGGG GCGCGGGCAG CCGCGCTCGG GCAGCGGGGG CCGGGGCTG 420  

CCGCTGCGC TCGCAGCTGG TGCCGCTGCG CGCGCTCGGC CTGGGCCACC GCTCCGACGA 480  

GCTGGTGCCT TTCCGCTTCT GCAGCGGCTC CTGCGCCGC GCGCGCTCTC CACACGACCT 540  

CAGCCTGGCC AGCCTACTGG GCGCCGGGGC CCTGCGACCG CCGCCGGGCT CCGGCCCGT 600  

CAGCCAGCCC TGCTGCCGAC CCACGCGCTA CGAAGCGGTC TCCTTCATGG ACGTCAACAG 660
  
```



CACCTGGAGA ACCGTGGACC GCCTCTCCGC CACCGCCTGC GGCTGCCTGG GCTGAGGGCT 720  
 CGCTCCAGGG CTTTGACAGC TGGACCCTTA CCGGTGGCTC TTCCTGCCTG GGACCCTCCC 780  
 GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGACGAAG GCCTCAAAGC TGAGAGGGCCC 840  
 CTACCGGTGG GTGATGGATA TCATCCCCGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900  
 CAGAGCCCTC ACCCTGCGGA TCCCAGCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960  
 CTTGCGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG GGACCCTCC 1020  
 TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT AGGGACAGCA 1080  
 TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAAGTGCC TGTGCTGGAA CTGGCCTGTA 1140  
 CTCACCTCATG GGAGCTGGCC CC

Seq ID NO: 281 Protein sequence:  
 Protein Accession #: NP\_476431.1

1 11 21 31 41 51  
 | | | | | |  
 MELGLGLST LSHCPWPRRQ APLGLSAQPA LWPTLAALAL LSSVAEASLG SAPRSPAPRE 60  
 GPPPVILASPA GHLPGGRTAR WCSGRARRPP PQPSRPAPPP PAPPSALPRG GRAARAGGPG 120  
 SRARAAGARG CRLRSQLVFV RALGLGHRSD ELVRFRCFSG SCRRARSPHD LSLASLLGAG 180  
 ALRPPPGSRP VSQPCCRPTR YEAVSFMDVN STWRTVDRLS ATACGCLG

Seq ID NO: 282 DNA sequence  
 Nucleic Acid Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | | |  
 CTACTGCACC TGCCCTCTGT TTCCTTTGGA AATCTCTTAC CTTTCATTAG GGTTTCTTTC 60  
 ATAGCAATTT CCTTTGGTTT TTAAGACTTC TACATTGCTT TTTCTTTTAT TATCTGTGCT 120  
 CCGTGAACCT TATGAATGCT GCTTAAAAAT AATGTCAAAA TATGTTTTAG CTGCCTACTC 180  
 AGGTAAACGTT TTCTTTTGCT CTCATCTTGG TTTCCATATA CTATTTTTGG TTTTTTGTGA 240  
 GATCTAATCA ATGATCTAGT CAGAAGCTAC TTCACTGGCT AACAGTGATC ATGTTTCATGT 300  
 GCTAAAAATG AACTTGAAAC ACGGAAGTAG TGTTTGGTCC AGTTTGAAAG CTCTTATTAG 360  
 TATTCTTCAT CCTGGCTGTA ATAATAGCCA TTATTTGTTA TGCCTTTGTG ATGTAGCAGA 420  
 CACTCTTAAG GATTTTATGT GTATTATTCA AATTGCTATT ACTGTTCTTT TTATAGTTGA 480  
 GAATCTCAGG ATACCTACAT TTATCACTTT TTCAATATAT ATGTATTTCT TATT

Seq ID NO: 283 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 564-1481

1 11 21 31 41 51  
 | | | | | |  
 GAGACTTTTA ATCATCTATC CCTTGTGCTT TACGCAGACC CTACAATACA CTAGAGGCTT 60  
 CAAAGAGGTC AAAAATTCAC ATGTGTAGAC AAATTAGGTC CCTTAAGATG CCAGGCAAAC 120  
 GAAGTGCTAC CAAAACACGC AATGATGTC CTAAAAGTGC GTTCTGGGAT ACACCTGTAA 180  
 ACTTGGATCA AGTTCCCTCC CCTCTCCTCA AAATAATATCG ACTTGTGCTG AAAGAAATCA 240  
 CGACCGATGC TCACAATTCT GACCTCGTAA TTATATAGGG GGTGGTTTTG GTTTCTGCGT 300  
 CTTTCCCTGA TTCAGTGGCA GGTAAACATAT TTCATGTACA AAATGAACTG CAACACCACG 360  
 GCAAACAAGG GACAGGCCCT CAAAGTTGTC GGTAGGAGC CAGGACCCCG CCAGTGGCGT 420  
 GGGGAGACAC CGTACTAAAC AAGCTTGCAA ACAGCAGGCA CCTTCCTGCC ACTGAGGAGG 480  
 AAGGGCTGGC TAAGGGAGGC CGGGCGGAGG GAAGCCAAGC TCTGCAGGCC CTGACAAAGT 540  
 CCTCCCGGCC TCCACGCGTC GGCATGGCAA CGCGGGTCT GTGCTGGCCG GGATTGGCCG 600  
 GCCTGGCGCG CGCAGGGCCC GCTGGGAAAG CGCGTCCCG CCGCGGCTCC GCCAGTTTGA 660  
 ACTTGGCGGG CCAGATGTGG GCGGCGGGGC GCTGGGGGCC TACTTTTCCC TCTTCCTACG 720  
 CCGGTTTCTC TGCTGACTGC AGACCCAGGT CTCGGCCCTC CTCGGACTCC TGCTCAGTCC 780  
 CTATGACGGG CGCAGCTGGG CAGGGGCTGG AGGTGGTGCG CTCGCCGTG CCGCCGCTGC 840  
 CGCTGAGCTG CAGCAATTCC ACCAGGTCGC TGTGTCTCC CCTTGGCCAC CAGAGCTTCC 900  
 AGTTTGACGA GGACGACGGT GACGGGGAGG ATGAGGAAGA CGTGGATGAT GAGGAAGACG 960  
 TGGATGAAGA TGCCCATGAT TCAGAGGCCA AAGTGGCGAG CCTGAGAGGA ATGGAGTTAC 1020  
 AGGGGTGCGC CAGCACTCAG GTTGAATCAG AAAATAACCA AGAAGAACAG AAACAGGTGC 1080  
 GCTTACCAGA AAGCCGCTG ACACCATGGG AGGTGTGGTT TATTGGCAA GAAAAAGAA 1140  
 AACGTGACCG GCTGCAACTG AAAGCTCTAG AGGAATTTAA TCAACAATA GAAAAAGAA 1200  
 AAGAAATGGA AGAACGTGAA AAAAGAAAGA TAATTGCTGA AGAAAAGCAC AAGGAATGGG 1260  
 TTCAGAAAAA GAATGAGCAA AAAAGAAAAG AAAGAGAACA AAAAAATTAAT AAAGAAATGG 1320  
 AGGAAAAAGC AGCAAAAGAA CTGGAGAAAG AATACTTGCA AGAAAAAGCA AAAGAAAAAT 1380  
 ATCAAGAATG GTTAAAGAAA AAAAATGCTG AAGAATGTGA GAGGAAGAAG AAAGAAAAGA 1440  
 AAAACAACAG CAAGCTGAAA TACAGGAGAA AAAGGAATA GCAGAAAAAA AGTTTCAAGA 1500  
 ATGGTTGGAA AATGCGAAAC ATAAACCTCG TCCAGCTGCA AAGAGCTATG GTTATGCCAA 1560  
 TGGAAACTT ACAGGTTTTT ACAGTGGAAT TTCCTATCCA GAACAGCCT TTTATAATCC 1620  
 AATTCCGTGG AAACCAATTC ATATGCCACC TCCCAAAGAA GCTAAGGATC TATCAGGAAG 1680  
 GAAGAGTAAA AGACCTGTGA TAAGTCAGCC ACACAAGTCA TCATCTCTGG TAATTCATAA 1740  
 AGCCAGGAGC AATCTTTGCC TTGGAACCTC GTGCAGAATA CAAAGATAGC GTATGTGGAA 1800  
 AATAACATGC TTTTATCTGG AGCTATTTAA TTTAAAAATC AGAAATTGTT TTTTACTGCT 1860  
 CAGTCAATAA CTCAACACTT AATGTGATTA TTGACAAATA GCAATTTTTG CATTTGTATA 1920  
 TGGAGTCCTT AGAGTTGAGG AAGATATTTT CTGGATTTTG GTTTTTATAA ACTTTTTAAG 1980  
 GTTGATCTTG GCATGTTGTT TTGCAGAATA AGTGGCTGAA TATGTAAGAA TTGTGTTTGT 2040  
 ATTTAGCTTG TATTAAAAGT AACTGTAAAT ACCAATAAAA CTAACAATTT TTCTTG

Seq ID NO: 284 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | | |  
 MATRGLCWPG LAGLARAGPA GKARPRRGS SLNLAGQMW AGRWGPTFPS SYAGFSADCR 60  
 PRSRPSSDSC SVPMTGARGQ GLEVVRSPSP PLPLSCSNST RSLLSPLGHQ SFQFDEDDGD 120  
 GEDEEDVDDE EDVDEDAHDS EAKVASLRGM ELQGCASTQV ESENNQBEQK QVRLPESRLT 180  
 PWEVWFIGKE KEERDRLQLK ALEELNQLE KRKEMEEREK RKIIAEEKHK EWVQKNEQK 240  
 RKEREQKINK EMEEKAAKEL EKEYLQEKAK EKYQEWLKKK NAECEKRRKK EKKNNSKLKY 300



RRKRRK

Seq ID NO: 285 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1-1746

5

1	11	21	31	41	51	
ATGCCACTGA	AGCATTATCT	CCTTTTGCTG	GTGGGCTGCC	AAGCCTGGGG	TGCAGGGTTG	60
GCCTACCATG	GCTGCCCTAG	CGAGTGTAAC	TGCTCCAGGG	CCTCCCAGGT	GGAGTGCACC	120
GGGGCACGCA	TTGTGGCGGT	GCCCCCCCT	CTGCCCTGGA	ACGCCATGAG	CCTGCAGATC	180
CTCAACACGC	ACATCACTGA	ACTCAATGAG	TCCCCGTTCC	TCAATATCTC	AGCCCTCATC	240
GCCCTGAGGA	TTGAGAAGAA	TGAGCTGTCT	CGCATCACGC	CTGGGGCCTT	CCGAAACCTG	300
GGCTCGCTGC	GCTATCTCAG	CCTCGCCAAC	AACAAGCTGC	AGGTTCTGCC	CATCGGCCTC	360
TTCCAGGGCC	TGGACAGCCT	TGAGTCTCTC	CTTCTGTCCA	GTAACCAGCT	GTTGCAGATC	420
CAGCCGGCCC	ACTTCTCCCA	GTGCAGCAAC	CTCAAGGAGC	TGCAGTTGCA	CGGCAACCAC	480
CTGGAATACA	TCCCTGACGG	AGCCTTCGAC	CACCTGGTAG	GACTCACGAA	GCTCAATCTG	540
GGCAAGAATA	GCCTCACCCA	CATCTCACCC	AGGGTCTTCC	AGCACCTGGG	CAATCTCCAG	600
GTCCTCCGGC	TGTATGAGAA	CAGGCTCACG	GATATCCCCA	TGGGCACTTT	TGATGGGCTT	660
GTTAACCTGC	AGGAACTGGC	TCTACAGCAG	AACCAGATTG	GACTGCTCTC	CCCTGGTCTC	720
TTCCACAACA	ACCACAACCT	CCAGAGACTC	TACCTGTCCA	ACAACCACAT	CTCCCAGCTG	780
CCACCCGACA	TCTTCATGCA	GCTGCCCCAG	CTCAACCGTC	TTACTCTCTT	TGGGAATTCC	840
CTGAAGGAGC	TCTCTCTGGG	GATCTTCGGG	CCCATGCCCA	ACCTGCGGGA	GCTTTGGCTC	900
TATGACAACC	ACATCTCTTC	TCTACCCGAC	AATGTCTTCA	GCAACCTCCG	CCAGTTGCAG	960
GTCCTGATTC	TTAGCCGCAA	TCAGATCAGC	TTTCTCTCCC	CGGGTGCCCT	CAACGGGCTA	1020
ACGGAGCTTC	GGGAGCTGTC	CCTCCACACC	AACGCACCTG	AGGACCTGGA	CGGGAATGTC	1080
TTCCGCATGT	TGGCCAACCT	GCAGAACATC	TCCCTGCAGA	ACAATCGCCT	CAGACAGCTC	1140
CCAGGGAATA	TCTTCGCCAA	CGTCAATGGC	CTCATGGCCA	TCCAGCTGCA	GAACAAACAG	1200
CTGGAGAACT	TGCCCCCTCG	CATCTTCGAT	CACCTGGGGA	AACTGTGTGA	GCTGCGGCTG	1260
TATGACAATC	CCTGGAGGTG	TGACTCAGAC	ATCCTTCCGC	TCCGCAACTG	GCTCCTGCTC	1320
AACCAGCCTA	GTTTAGGGAC	GGACACTGTA	CCTGTGTGTT	TCAGCCCAGC	CAATGTCCGA	1380
GGCCAGTCCC	TCATTATCAT	CAATGTCAAC	GTGTCTGTTC	CAAGCGTCCA	TGTCCCTGAG	1440
GTGCCTAGTT	ACCCAGAAAC	ACCATGGTAC	CCAGACACAC	CCAGTTACCC	TGACACCACA	1500
TCCGTCTCTT	CTACCACTGA	GCTAACACAG	CCTGTGGAAG	ACTACACTGA	TCTGACTACC	1560
ATTCAGGTCA	CTGATGACCG	CAGCGTTTGG	GGCATGACCC	AGGCCAGAG	CGGGCTGGCC	1620
ATTGCCGCCA	TTGTAATTGG	CATTGTGCGC	CTGGCCTGCT	CCCTGGCTGC	CTGCGTCGGC	1680
TGTTGCTGCT	GCAAGAAGAG	GAGCCAAGCT	GTCTGTATGC	AGATGAAGGC	ACCCAATGAG	1740
TGTTAAAGAG	GCAGGCTGGA	GCAGGGCTGG	GGATGATGG	GACTGGAGGA	CCTGGGAATT	1800
TCATCTTTCT	GCCTCCACCC	CTGGGTCCAT	GGAGCTTTCC	CGTGATTGCT	CTTTCTGGCC	1860
CTAGATAAAG	GTGTGCCTAC	CTCTTCCTGA	CTTGCTGTAT	TCTCCCGTAG	AGAAGCAGGT	1920
CGTGCCGGAC	CTTCCTACAA	TCAGGAAGAT	AGATCCAAC	GGCCATGGCA	AAAGCCCTGG	1980
GGATTTCGGA	TTCATACCCC	TGGGCTTCCT	TCGAGAGGGC	TCTTCCTCCA	AATCCTCCCC	2040
ACCTGTCTCT	CAAGAACAGC	CTTCCCTGCG	CCCAGGCCCC	CTCCGGGCTT	CTGTAGACTC	2100
AGTTAGTCCA	CAGCTGCTC	ACTTCGTGGG	AATAGTTCTC	CGCTGAGATA	GCCCCCTCTG	2160
CCTAAGTATT	ATGTAAGTTG	ATTTCCCTTC	TTTTGTTTCT	CTTGTTTGTG	CTATGGCTTG	2220
ACCCAGCATG	TCCCTCAAA	TGAAAGTTCT	CCCCTTGATT	TTCTGCTCCT	GAAGGCAGGG	2280
TGAGTTCTCT	CCTCAAAGAA	GACTTCAACT	CATTTAAGTG	GTTTCTTAAG	AGCCGTCAAT	2340
CAGCCTGGTT	TTGGGGATGC	TATGAAAGAG	AGAAGGAAAA	TCATGCCGCT	CAGTTCCTGG	2400
AGACAGAAGA	GCCGTCATCA	GTGTCTCACT	TGTGATTTTT	ATCTGAAAA	GGAAGAAACA	2460
CCCCAGCACA	GCAAGCTCAG	CCTTTTAGAG	AAGGATATTT	CCAACTGCA	AACTTTGCTT	2520
TGAAAAGTTT	AGCCCTTTAA	GGAATGAAAT	CATGTAGAAT	TTTGACTTTC	TAAAAACATT	2580
AAAATCAGCT	TATTAATACG	GAATAGAGAA	AGAAATCTGG	TGCCTGGGGG	TCCCTGTGTT	2640
CACCCCTAGA	GTTTGTTTTA	AAATTTTAA	TTGAAGCATG	TGAAGTGATC	STGCAGAAAA	2700
GTGGGAACAT	GATAGTGAT	GGCTTGGTGG	ATTTTCACAA	ACTGAACATA	CCTGTGTAAT	2760
CAGCATCTAG	ACCCAGACCC	AGAGCATCAC	AAATATCCCC	CATCCTGGGC	TTTTCCCGAG	2820
GGAGATGGGG	GCTTCTGAAG	ATGGACTTAC	CTGGGACCTG	CCCCCATGA	GCCAGGACGG	2880
TCCCCCACA	GTCAGCCTGT	GCAAAGGCCC	CGTGGCCAGG	GGTGGAGGAG	AATATGTGGG	2940
TGTGGACAGG	ATGGGAGACT	GTGGCCTGAA	CAGGAGATTT	TATTATATCT	GGAGACCCTG	3000
AGAGACCCTG	AGACCTGGGG	CACCATGGTG	GGCCAGGTCA	GAAGCATCCT	GACTGCAGAG	3060
GTCCGTGCAG	CCACACCCCT	TTCCCTGCCA	GCAAGTTGTC	TGCGGCTCAT	CGGAGGCCCC	3120
TCCGCCTGGA	GCCTTCTATG	GACGTGATAT	GCCTGTATCT	GTTTTTAATT	TTCATTCTTC	3180
ACTTAGGGGA	AGTGAAATCG	CTCAGAGATG	AGATCCTTTA	ATTGAAAACG	AAGTGTAACG	3240
GAATCTAGTG	TCTTTCTAAT	GTGGTAAAA	TCTCCATCAA	CATCACAGTC	AGCTGGCAGC	3300
TGAACCTCAG	AATCTCACTT	ACAGCAGGCG	ACACGGGGGT	ACACCGATGG	GTCACACTGG	3360
GTCTGGGGGC	TCCCTGGAGC	TCCTCCTGCG	TGTGGTCTGG	TTAGGAGTTG	AGTTGTTTGC	3420
TCCAGGGTTA	TTCTCCTCCT	CGAGTCAAG	TCACACGAAT	ACCTGCCTTC	TCTGGCTTTC	3480
CTGCTATACA	CATATTCACA	TGGCGCTCAA	GAAGTTAGGC	TCATGGCAAC	GTGTGTCTTT	3540
CTCTGGACAA	CTGGCCAGT	TTACAGTGAA	ATGGAGAATT	TCAGGTCTCC	ACGTCTGCCC	3600
AGGAAAGAAC	TTCAGCTGAC	TCCACGGGGA	TCTGGAAATC	CACGACCAAT	CCCGATCGGC	3660
TCTTATTAGC	TCCCCGCTCC	ACAAGACACC	TGTGCTTTGG	AAATCCACCA	CCAATCCCGA	3720
TCCGGCTCTTA	TTAGCTCCCC	GCTCCACAAG	ACACCTGTGA	TCTGGAAATC	TACCACCAAT	3780
CCCGATCGGC	TCTTATTAGC	TCCCCGCTCC	ACAAGACACC	TGTGACATCC	TCCAGGGCCA	3840
CAGGAGCACG	TGCTGACCAG	TTTTCCCTTC	CAGTTCCTGC	ACAAAAAGTG	TCCAGAGGGC	3900
TGTTTGCAAA	CACTAGTGCA	CTTTGTAGCT	TTTCACCCTC	TGTCCCAGGG	AATCTAGGAG	3960
AGATGAGGCC	CGTCAGAGTC	AAGAGATGTC	ATCCCCCAG	GGTCTCCAAG	GCATTTCCAC	4020
ACTATTGGTG	GCACCTGGAG	GACATGCACC	AAGGCTTGCC	AGAGCCAACA	GGAAGTGAGC	4080
CCAGAGCATG	GCACATGAGC	ATCACCCGCT	GATGGTGGCC	TGCTGTGCCT	GGTGCCAACA	4140
GGGGCATCCC	GGCCCGTACC	CCTCCAGACA	GGAAGCATGG	GTTTGCCAC	AGACCTGTCT	4200
GGTGCTCCTG	TGAGTGGCCT	CCAGATGTCT	TTGTGCATAG	GCACAAGTGG	GCCAGGGCTG	4260
GAGGGAGGTG	GGAAACCTCA	TCATCCGGTG	GGCCCTGCCA	ATCTTAACCC	AGAACCCTTA	4320
GGTATTCCTG	GCAGTAGCCA	TGACATTGGA	GCACCTTCCT	CTCCAGCCAG	AGGCTGACCT	4380
GAGGGCCACT	GTCCTCAGAT	GACACCACCC	AGGAGCACCC	TAGGTGAGGG	GTGAGGGCCC	4440
CCTTATGTGA	ACCTCTTGCC	TCTTCTTTTC	TCCCATCAGA	GTGGTTGGAT	GGAGCCATTG	4500
GCCTCCTTTT	CTTCAGCGGG	CCCTTCAACC	TCTCTGCACC	ATGTTGTCTG	GCTGAGGAGC	4560
TACTAGAAAA	GCTGAGTGGG	GCTTCTTTTC	CAACAGGATG	ATGCATTTGC	TCAATTCTCA	4620
GGGCTGGAAT	GAGCCGGCTG	GTCCCCCAGA	AAGCTGGAGT	GGGGTACAGA	GTTCACTTTT	4680
CCTCTCTGTT	TACAGCTCCT	TGACAGTCCC	ACGCCCATCT	GGAGTGGGAG	CTGGGAGTTA	4740



```

GTGTTGGAGA AGAAACAACA AAAGCCAATT AGAACCCTA TTTTAAAAA GTGCTTACTG 4800
TGCACAGATA CTCTTCAAGC ACTGGACGTG GATTCTCTCT CTAGCCCTCA GCACCCCTGC 4860
GGTAGGAGTG CCGCCTCTAC CCACTTGTGA TGGGGTACAG AGGCACTTGC TCTTCTGCAT 4920
GGTGTTCAT AGGCTGGGAG TTTTATTTAT CTCTTCAAC TTTGTACAAG AGCTCATGGC 4980
5 TGTCTTGGG CTTTCGTCAT TAAACCAAAG GAAATGGAAG CCATTCCTCT GTTGCTCTCC 5040
TTAGTCTTGG TCATCAGAAC CTCACCTGGT ACCATATAGA TCAAAAGCTT TGTAAACCACA 5100
GGAAAAATA AACTCTTCCA TCCCTTAAAG AATAGAATAG TTTGTCCCTC TCATGGGAAT 5160
TGGGCTGTAT GTATATTGTT CTTCTCTCTT AGAATTTAGA GATACAAGAG TTCTACTTAG 5220
10 AACTTTTCAT GGACACAATT TCCACAACCT TTCAGATGCT GATGTAGAGC TATTGGGAAA 5280
GAACTTCCAA ACTCAGGAAG TTTGCAGAGA GCAGACAGCT AGAGATAACT CGGGACCCAG 5340
AGTTGGTCGA CAGATGTTAG ATGTATCCTA GCTTTTAGCC ATAAACCACT CAAAGATTCA 5400
GCCCCCAGAT CCCACAGTCA GAACTGAATC TGCGTTGTTG GGAAGCCAGC AGTGGCCTTG 5460
GGAAGGAAGC CATGGCTGTG GTTCAGAGAG GGTGGGCTGG CAAGCCACTT CCGGGGAAAA 5520
15 CTCCTTCCGC CCCAGGTTTC TTCTTCTCTT AAGGAGAGAT TGTTCCTACC AATCCGCTGC 5580
CTTCATGCTG CTTTCAAAGC TAGATCATGT TTGCCTTGCT TAGAGAATTA CTGCAATCA 5640
GCCCCAGTGC TTGGCGATGC ATTTACAGAT TTCTAGGCCC TCAGGGTTTT GTAGAGTGTG 5700
AGCCCTGGTG GGCAGGGTTG GGGGGTCTGT CTTCTGCTGG ATGCTGCTTG TAATCCATTT 5760
GGTGTACAGA ATCAACAATA AATAATATAC ATGTAT

```

Seq ID NO: 286 Protein sequence:  
Protein Accession #: NP\_570843.1

```

1      11      21      31      41      51
|      |      |      |      |      |
25 MPLKHYLLLL VGCQAWGAGL AYHGCPSECT CSRASQVECT GARIVAVPTP LPWNAMSLQI 60
LNTHITELNE SPFLNISALI ALRIEKNELS RITPGAFRNL GSLRYLSLAN NKLVLPPIGL 120
FQGLDSLESL LLSSNQLLQI QPAHFSQCSN LKELQLHGNH LEYIPDGAFD HLVGLTKLNL 180
GKNSLTHISP RVFQHLGNLQ VLRLYENRLT DIPMGTFDGL VNLQELALQQ NQIGLLSPGL 240
30 FHNHNLQRL YLSNNHISQL PPSIFMQLPQ LNRLTLFGNS LKELSLGIFG PMPNLRELWL 300
YDNHISLPD NVFSNLRQLQ VLILSRNQIS FISPGAFNGL TELRELSLHT NALQDLGDNV 360
FRMLANLQNI SLQNNRLRQL PGNIFANVNG LMAIQLQNNQ LENLPLGIFD HLGKLCLERL 420
YDNPWRCDSD ILPLRNWLLL NQPRLGTDV PVCFSANVR GQSLIIINVN VAVPSVHVPE 480
VPSYPETPWY PDTPSYPDIT SVSSTTELT S PVEDYTDLT IQVTDDRSVW GMTQAQSGLA 540
35 IAAIVIGIVA LACSLAACVG CCCCKKRSQA VLMQMKAPNE C

```

Seq ID NO: 287 DNA sequence  
Nucleic Acid Accession #: NM\_002362  
Coding sequence: 1..954

```

1      11      21      31      41      51
|      |      |      |      |      |
40 ATGTCTTCTG AGCAGAAAGAG TCAGCACTGC AAGCCTGAGG AAGGCGTTGA GGCCCAAGAA 60
GAGGCCCTGG GCCTGGTGGG TGCACAGGCT CCTACTACTG AGGAGCAGGA GGCTGCTGTC 120
45 TCCTCCTCCT CTCTCTGCTT CCCTGGCACC CTGGAGGAAG TGCTGCTGTC TGAGTCAGCA 180
GGTCTCCTCC AGAGTCTCTA GGGAGCCTCT GCCTTACCCA CTACCATCAG CTTCACTTGC 240
TGGAGGCAAC CCAATGAGGG TTCCAGCAGC CAAGAAGAGG AGGGGCCAAG CACCTCGCCT 300
GACGCAGAGT CCTTGTTCCT AGAAGCACTC AGTAACAAGG TGGATGAGTT GGCTCATTTT 360
CTGCTCCGCA AGTATCGAGC CAAGGAGCTG GTCACAAAGG CAGAAATGCT GGAGAGAGTC 420
50 ATCAAAAATT ACAAGCGCTG CTTTCTCTGT ATCTTCGCA AAGCCTCCGA GTCCCTGAAG 480
ATGATCTTTG GCATTGACGT GAAGGAAGTG GACCCCGCCA GCAACACCTA CACCCTTGTC 540
ACCTGCCTGG GCCTTTCCTA TGATGGCCTG CTGGGTAATA ATCAGATCTT TCCCAAGACA 600
GGCCTTCTGA TAATCGTCTT GGGCACAATT GCAATGGAGG GCGACAGCGC CTCTGAGGAG 660
GAAATCTGGG AGGAGCTGGG TGTGATGATG GGAGGGAGCA CACTGTCTAT 720
55 GGGGAGCCCA GGAAACTGCT CACCCAAGAT TGGGTGCAGG AAAACTACCT GGAGTACCGG 780
CAGGTACCCG GCAGTAATCC TGCGCGCTAT GAGTTCCTGT GGGGTCCAAG GGCTCTGGCT 840
GAAACCAGCT ATGTGAAAGT CCTGGAGCAT GTGGTCAGGG TCAATGCAAG AGTTCGCATT 900
GCCTACCCAT CCCTGCGTGA AGCAGCTTTG TTAGAGGAGG AAGAGGGAGT CTGA

```

Seq ID NO: 288 Protein sequence:  
Protein Accession #: NP\_002353.1

```

1      11      21      31      41      51
|      |      |      |      |      |
65 MSSEQKSQHC KPEGVEAQE EALGLVGAQA PTTEEQEAIV SSSSPLVPPT LEEVPAESA 60
GPPQSPQGAS ALPTTISFTC WRQPNEGSSS QEEEGPSTSP DAESLFREAL SNKVDELAHF 120
LLRKYRAKEL VTKAEMLERV IKNYKRCFPV IFGKASESLK MIFGIDVKEV DPASNTYTLV 180
TCLGLSYDGL LGNNQIFPKT GLLIIVLGTI AMEGDSASEE EIWEELGVMG VYDGREHTVY 240
70 GEPRKLLTQD WQENYLEYR QVPGSNPARY EFLWGPRLA ETSYVKVLEH VVRVNARVRI 300
AYPSLREAAAL LEEEEGV

```

Seq ID NO: 289 DNA sequence  
Nucleic Acid Accession #: NM\_002362  
Coding sequence: 46..1344

```

1      11      21      31      41      51
|      |      |      |      |      |
80 CGGCGGCCGC GCCCTGGTTG GGTCCCCACT GCTCTCGGGG GCGCCATGGA CGAGGCCGTG 60
GGCGACCTGA AGCAGGCGCT TCCCTGTGTG GCGAGTCGCG CAACGGTCCA CGTGGAGGTG 120
CATCAGCGCG GCAGCAGCAC TGCAAAGAAA GAAGACATAA ACCTGAGTGT TAGAAAGCTA 180
CTCAACAGAC ATAATATTGT GTTTGGTGAT TACACATGGA CTGAGTTTGA TGAACCTTTT 240
TTGACCAGAA ATGTGCAGTC TGTGTCTATT ATTGACACAG AATTAAAGGT TAAAGACTCA 300
CAGCCCATCG ATTTGAGTGC ATGCTACTGT GCACTTCACA TTTTCCAGCT GAATGAAGAT 360
85 GGCCCCAGCA GTGAAATCTT GGAGGAAGAG ACAGAAAACA TAATTGCAGC AAATCACTGG 420
GTTCTACCTG CAGCTGAATT CCATGGGCTT TGGGACAGCT TGGTATACGA TGTGGAAGTC 480
AAATCCCATC TCCTCGATTA TGTGATGACA ACTTTACTGT TTTGAGACAA GAACGTCAAC 540

```



5  
10  
15  
20  
25  
30

AGCAACCTCA	TCACCTGGAA	CCGGGTGGTG	CTGCTCCACG	GTCCTCCTGG	CACTGGAAAA	600
ACATCCCCTGT	GTAAAGCGTT	AGCCCAGAAA	TTGACAATTA	GACTTTCAAG	CAGGTACCGA	660
TATGGCCCAAT	TAATTGAAAT	AAACAGCCAC	AGCCTCTTTT	CTAAGTGGTT	TTCGGAAAGT	720
GGCAAGCTGG	TAACCAAGAT	GTTTCAGAAG	ATTGAGGATT	TGATTGATGA	TAAAGACGCC	780
CTGGTGTTCG	TGCTGATTGA	TGAGGTGGAG	AGTCTCACAG	CCGCCCCGAAA	TGCCTGCAGG	840
GCGGGCACCG	AGCCATCAGA	TGCCATCCGC	GTGGTCAATG	CTGTCTTGAC	CCAAATTGAT	900
CAGATTAAAA	GGCATTCCAA	TGTTGTGATT	CTGACCACTT	CTAACATCAC	CGAGAAGATC	960
GACGTGGCCT	TCGTGGACAG	GGCTGACATC	AAGCAGTACA	TTGGGCCACC	CTCTGCAGCA	1020
GCCATCTTCA	AAATCTACCT	CTCTTGTTTG	GAAGAAGTGA	TGAAGTGTC	GATCATATAC	1080
CCTCGCCAGC	AGCTGCTGAC	CCTCCGAGAG	CTAGAGATGA	TTGGCTTCAT	TGAAAACAAC	1140
GTGTCAAAAT	TGAGCCTTCT	TTTGAATGAC	ATTTCAAGGA	AGAGCGAGGG	CCTCAGCGGC	1200
CGGGTCCCTGA	GAAAACTCCC	CTTTCTGGCT	CATGCGCTGT	ATGTCCAGGC	CCCCACCGTC	1260
ACCATAGAGG	GGTTCCTCCA	GGCCCTGTCT	CTGGCAGTGG	ACAAGCAGTT	TGAAGAGAGA	1320
AAGAAGCTTG	CAGCTTACAT	CTGATCCTGG	GCTTCCCCAT	CTGGTGCTTT	TCCCATGGAG	1380
AAACACAAAC	CAGTAAGTGA	GGTTGCCCTC	CACAGCCGTC	TCCAGGGGAA	TCCCTTCTGC	1440
AAACCAAACG	TTACTTAGAC	TGCAAGCTAG	AAAGCCACCA	AGGCCAGGCT	TTGTTAAAAG	1500
AAGTGTATTC	TATTTATGTT	GTTTTAAAAT	GCATACTGAG	AGACAAACAT	CTTGTCATTT	1560
TCACTGTTTG	TAAAAGATAA	TTCAGATTGT	TTGTCTCCTT	GTGAAGAACC	ATCGAAACCT	1620
GTTTGTTCCC	AGCCCCCCCC	CAGTGGATGG	GATGCATAAT	GCCAGCAAGT	TTTGTTTAAC	1680
AGCAAAAAAG	GAAGATTAAT	GCAGGTGTTA	TAGAAGCCAG	AAGAGAAACT	GTGTCACCTT	1740
AAAGAAGCAT	ATAATCATAG	CATTAATAAT	GCACACATTA	CTCCAGGTGG	AAGGTGGCAA	1800
TTGCTTTCTG	ATATCAGCTC	GTTTGATTTA	GTGCAAAAAT	GTTTTCAAGA	CTATTTAATG	1860
GATGTAAAAA	AGCCTATTTT	TACATTATAC	CAACTGAGAA	AAAAATGGTC	GGTAAAGTGT	1920
TCTTTTATAA	TAAATAATCA	AGACATGGTC	CCATTTGCAG	GAAAAGTGCA	GACTCTGAGT	1980
GTTCCAGGGA	AACACATGCT	GGACATCCCT	TGTAACCCGG	TATGGGCGCC	CCTGCATTGC	2040
TGGGATGTTT	CTGCCACCGG	TTTTGTTTGT	GCAATAACGT	TATCACATTT	CTAATGAGGA	2100
TTCACATTAA	TATAATATAA	AATAAATAGG	TCAGTTACTG	GTCTCTTTCT	GCCGAATGTT	2160
ATGTTTTGCT	TTTATCTCAC	AGTAAAATAA	ATATAATTAA	AAA		

Seq ID NO: 290 Protein sequence:  
Protein Accession #: NP\_004228

35  
40

1	11	21	31	41	51	
MDEAVGDLKQ	ALPCVAESPT	VHVEVHQRGS	STAKKEDINL	SVRKLNRHN	IVFGDYTWTE	60
FDEPFLTRNV	QSVSIIDTEL	KVKDSQPIDL	SACTVALHIF	QLNEDGPSSE	NLEEETENII	120
AANHWWLPAA	EFHGLWDSL	YDVEVKSHLL	DYVMTLLFS	DKNVNSNLIT	WNRVVLHGP	180
PGTGKTSCLK	ALAQLTIRL	SSRYRYQQLI	EINSHSLFSK	WFSESGKLVT	KMFQKIQDLI	240
DDKDALVFVL	IDEVESLTAA	RNACRAGTEP	SDAIRVNVAV	LTQIDQIKRH	SNVVILTTSN	300
ITEKIDVAFV	DRADIKQYIG	PPSAAAFKFI	YLSCLEELMK	CQIIYPRQQL	LTLRELEMIG	360
FIENNVSKLS	LLLNDISRKS	EGLSGRVLRL	LPFLAHALYV	QAPTVTIEGF	LQALSLAVDK	420
QFEERKKLAA	YI					

Seq ID NO: 291 DNA sequence  
Nucleic Acid Accession #: NM\_002658.1  
Coding sequence: 77-1372

50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
GTCCCCGCAG	CGCCGTCGCG	CCCTCCTGCC	GCAGGCCACC	GAGGCCGCCG	CCGTCTAGCG	60
CCCCGACCTC	GCCACCATGA	GAGCCCTGCT	GGCGCGCCTG	CTTCTCTGCG	TCCTGGTCTG	120
GAGCGACTCC	AAAGGCAGCA	ATGAACCTTA	TCAAGTTCCA	TCAAGCTGTG	ACTGTCTAAA	180
TGGAGGAACA	TGTGTGTCCA	ACAAGTACTT	CTCCAACATT	CAGTGTGCA	ACTGCCCAAA	240
GAAATTTCGA	GGGCAGCACT	GTGAAATAGA	TAAGTCAAAA	ACCTGCTATG	AGGGGAATGG	300
TCACCTTTAC	CGAGGAAAGG	CCAGCACTGA	CACCATGGGC	CGGCCCTGCC	TGCCCTGGAA	360
CTCTGCCACT	GTCCTTCAGC	AAACCTACCA	TGCCCCACAG	TCTGATGCTC	TTCAGCTGGG	420
CCTGGGGAAA	CATAATTACT	GCAGGAACCC	AGACAACCGG	AGGCGACCCCT	GGTGCTATGT	480
GCAGGTGGGC	CTAAAGCCGC	TTGTCCAAGA	GTGCATGGTG	CATGACTGCG	CAGATGGAAA	540
AAAGCCCTCC	TCTCCTCCAG	AAGAATTAAA	ATTTCAAGTG	GGCCAAAAGA	CTCTGAGGCC	600
CCGCTTTAAG	ATTATTGGGG	GAGAATTCAC	CACCATCGAG	AACCAGCCCT	GGTTTGCGGC	660
CATCTACAGG	AGGCACCGGG	GGGGCTCTGT	CACCTACGTG	TGTGGAGGCA	GCCTCATCAG	720
CCCTTGCTGG	GTGATCAGCG	CCACACACTG	CTTCATTGAT	TACCCAAAGA	AGGAGGACTA	780
CATCGTCTAC	CTGGGTCGCT	CAAGGCTTAA	CTCCAACACG	CAAGGGGAGA	TGAAGTTTGA	840
GGTGGAAAAC	CTCATCCTAC	ACAAGGACTA	CAGCGCTGAC	ACGCTTGCTC	ACCACAACGA	900
CATTGCCTTG	CTGAAGATCC	GTTCCAAGGA	GGCAGGTGT	GCGCAGCCAT	CCCGGACTAT	960
ACAGACCATC	TGCCTGCCCT	CGATGTATAA	CGATCCCCAG	TTTGGCACAA	GCTGTGAGAT	1020
CACTGGCTTT	GGAAAAGAGA	ATTCTACCGA	CTATCTCTAT	CCGGAGCAGC	TGAAAATGAC	1080
TGTTGTGAAG	CTGATTTCCC	ACCGGGAGTG	TCAGCAGCCC	CACACTACTG	GCTCTGAAGT	1140
CACCACCAAA	ATGCTATGTG	CTGCTGACCC	CCAATGGAAA	ACAGATTCCCT	GCCAGGGAGA	1200
CTCAGGGGGA	CCCCTCGTCT	GTTCCTCCA	AGGCCGCGATG	ACTTTGACTG	GAATTGTGAG	1260
CTGGGGCCGT	GGATGTGCCC	TGAAGGACAA	GCCAGGCGTC	TACACGAGAG	TCTCACACTT	1320
CTTACCCTGG	ATCCGCGAGT	ACACCAAGGA	AGAGAATGGC	CTGGCCCTCT	GAGGGTCCCC	1380
AGGGAGGAAA	CGGGCACCAC	CCGCTTTCTT	GCTGGTTGTC	ATTTTTGCAG	TAGAGTCATC	1440
TCCATCAGCT	GTAAGAAGAG	ACTGGGAAGA	TAGGCTCTGC	ACAGATGGAT	TTGCCTGTGG	1500
CACCACCAGG	GTGAACGACA	ATAGCTTTAC	CCTCACGGAT	AGGCCTGGGT	GCTGGCTGCC	1560
CAGACCTCTT	GGCCAGGATG	GAGGGGTGGT	CCTGACTCAA	CATGTTACTG	ACCAGCAACT	1620
TGTCTTTTTC	TGGACTGAAG	CCTGCAGGAG	TTAAAAAGGG	CAGGGCATCT	CCTGTGCATG	1680
GGTTCGAAGG	GAGAGCCAGC	TCCCCCGACC	TGTGAGGCCAT	TGTGAGGCC	ATGGTTGAGA	1740
AATGAATAAT	TTCCCAATTA	GGAAGTGTA	GCAGCTGAGG	TCTCTGAGG	GAGCTTAGCC	1800
AATGTGGGAG	CAGCGGTTTG	GGGAGCAGAG	ACACTAACGA	CTTCAGGGCA	GGGCTCTGAT	1860
ATTCCATGAA	TGTATCAGGA	AATATATATG	TGTGTGTATG	TTTGACACT	TGTTGTGTGG	1920
GCTGTGAGTG	TAAGTGTGAG	TAAGAGCTGG	TGTCTGATTG	TTAAGTCTAA	ATATTTCTTT	1980
AAACTGTGTG	GACTGTGATG	CCACACAGAG	TGGTCTTTCT	GGAGAGGTTA	TAGGTCATCT	2040
CTGGGGCCCT	TTGGGTCCCC	CACGTGACAG	TGCCTGGGAA	TGTACTTATT	CTGCAGCATG	2100
ACCTGTGACC	AGCACTGTCT	CAGTTTCACT	TTACACATAGA	TGTCCCTTTC	TTGGCCAGTT	2160
ATCCCTTCCT	TTTAGCCTAG	TTTATCCAA	CCTCACTGGG	TGGGGTGAGG	ACCACTCCTT	2220
ACACTGAATA	TTTATATTTT	ACTATTTTTT	TTTATATTTT	TGTAATTTTA	AATAAAAAGTG	2280



ATCAATAAAA TGTGATTTTT CTGA

Seq ID NO: 292 Protein sequence:  
Protein Accession #:NP\_002649.1

1	11	21	31	41	51	
MRALLARLLL	CVLVVSDSKG	SNELHQVPSN	CDCLNGGTCV	SNKYFSNIHW	CNCPKKFGGQ	60
HCEIDKSKTC	YEGNGHFYRG	KASTDTMGRP	CLPWN SATVL	QQTYHAHRSD	ALQLGLGKHN	120
YCRNPKNRRR	PWCYVQVGLK	PLVQECMVHD	CADGKKPSSP	PEELKFQCGQ	KTLRPRFKII	180
GGEFTTIENQ	PWFAAIYRRH	RGGSVTYVCG	GSLISPCWVI	SATHCFIDYP	KKEDYIVYLG	240
RSRLNSNTQG	EMKFEVENLI	LHKDYSADTL	AHHNDIALLK	IRSKEGRCAQ	PSRTIQITICL	300
PSMYNDPQFG	TSCEITGFGK	ENSTDYLYPE	QLKMTVVKLI	SHRECQOPHY	YGSEVTTKML	360
CAADPQWKTD	SCQGDSSGGL	VCSLQGRMTL	TGIVSWGRGC	ALKDKPGVYT	RVSHFLPWIR	420
SHTKEENGLA	L					

Seq ID NO: 293 DNA sequence  
Nucleic Acid Accession #: NM\_001498  
Coding sequence: 93..2006

1	11	21	31	41	51	
GGCAGGAGGC	TGAGTGTCCT	TCTCGCGCCC	GGAAGCGGGC	GACCGCCGTC	AGCCCGGAGG	60
AGGAGGAGGA	GGAGGAGGAG	GAGGGGGCGG	CCATGGGGCT	GCTGTCCAG	GGCTCGCCGC	120
TGAGCTGGGA	GGAAACCAAG	CGCCATGCCG	ACCACGTGCG	GCGGCACGGG	ATCCTCCAGT	180
TCCTGCACAT	CTACCACGCC	GTCAAGGACC	GGCACAAGGA	CGTTCTCAAG	TGGGGCGATG	240
AGGTGGAATA	CATGTTGGTA	TCTTTTGATC	ATGAAAATAA	AAAAGTCCGG	TTGGTCTCTGT	300
CTGGGGAGAA	AGTTCTTGAA	ACTCTGCAAG	AGAAGGGGGA	AAGGACAAAC	CCAAACCATC	360
CTACCCTTTG	GAGACCAGAG	TATGGGAGTT	ACATGATTGA	AGGGACACCA	GGACAGCCCT	420
ACGGAGGAAC	AATGTCCGAG	TTCAATACAG	TTGAGGCCAA	CATGCGAAAA	CGCCGGAAGG	480
AGGCTACTTC	TATATTAGAA	GAAAATCAGG	CTCTTTGCAC	AATAACTTCA	TTTCCCAGAT	540
TAGGCTGTCC	TGGGTTTACA	CTGCCCCGAG	TCAAACCCAA	CCCAGTGGAA	GGAGGAGCTT	600
CCAAGTCCCT	CTTCTTTCCA	GATGAAGCAA	TAAACAAGCA	CCCTCGCTTC	AGTACCTTAA	660
CAAGAAATAT	CCGACATAGG	AGAGGAGAAA	AGGTTGTCTAT	CAATGTACCA	ATATTTAAGG	720
ACAAGAATAC	ACCATCTCCA	TTTATAGAAA	CATTTACTGA	GGATGATGAA	GCTTCAAGGG	780
CTTCTAAGCC	GGATCATATT	TACATGGATG	CCATGGGATT	TGGAATGGGC	AATTGCTGTC	840
TCCAGGTGAC	ATTCCAAGCC	TGCAGTATAT	CTGAGGCCAG	ATACCTTTAT	GATCAGTTGG	900
CTACTATCTG	TCCAATTGTT	ATGGCTTTGA	GTGCTGCATC	TCCCTTTTAC	CGAGGCTATG	960
TGTCAGACAT	TGATTGTGCG	TGGGGAGTGA	TTTCTGCATC	TGTAGATGAT	AGAACTCGGG	1020
AGGAGCGAGG	ACTGGAGCCA	TTGAAGAACA	ATAACTATAG	GATCAGTAAA	TCCCAGATATG	1080
ACTCAATAGA	CAGCTATTTA	TCTAAGTGTG	GTGAGAAATA	TAATGACATC	GACTTGACGA	1140
TAGATAAAGA	GATCTACGAA	CAGCTGTTGC	AGGAAGGCAT	TGATCATCTC	CTGGCCCAGC	1200
ATGTTGCTCA	TCTCTTTATT	AGAGACCCAC	TGACACTGTT	TGAAGAGAAA	ATACACCTGG	1260
ATGATGCTAA	TGAGTCTGAC	CATTTTGAGA	ATATTCAGTC	CACAAATTGG	CAGACAATGA	1320
GATTTAAGCC	CCCTCCTCCA	AACTCAGACA	TTGGATGGAG	AGTAGAATTT	CGACCCATGG	1380
AGGTGCAATT	AACAGACTTT	GAGAACTCTG	CCTATGTGGT	GTTTGTGGTA	CTGCTCACCA	1440
GAGTGATCCT	TTCCTACAAA	TTGGATTTTC	TCATTCCACT	GTCAAAGGTT	GATGAGAACA	1500
TGAAGGTAGC	ACAGAAAAGA	GATGCTGTCT	TGCAGGGAAT	GTTTATTTTC	AGGAAAGATA	1560
TTTGCAAAGG	TGGCAATGCA	GTGGTGGATG	GTTGTGGCAA	GGCCCAGAAC	AGCACGGAGC	1620
TCGCTGCAGA	GGAGTACACC	CTCATGAGCA	TAGACACCAT	CATCAATGGG	AAGGAAGGTG	1680
TGTTTTCTGG	ACTGATCCCA	ATTCTGAAC	CTTACCTTGA	AAACATGGAA	GTGGATGTGG	1740
ACACCAGATG	TAGTATTCTG	AACTACCTAA	AGCTAATTAA	GAAGAGAGCA	TCTGGAGAAC	1800
TAATGACAGT	TGCCAGATGG	ATGAGGGAGT	TTATCGCAAA	CCATCCTGAC	TACAAGCAAG	1860
ACAGTGTGAT	AACTGATGAA	ATGAATTATA	GCCTTATTTT	GAAGTGTAAC	CAAATTGCAA	1920
ATGAATTATG	TGAATGCCCA	GAGTTACTTG	GATCAGCATT	TAGGAAAAGT	AAATATAGTG	1980
GAAGTAAAC	TGACTCATCC	AACTAGACAT	TCTACAGAAA	GAAAAATGCA	TTATTGACGA	2040
ACTGGCTACA	GTACCATGCC	TCTCAGCCCC	TGTGTATAAT	ATGAAGACCA	AATGATAGAA	2100
CTGTACTGTT	TTCTGGGCCA	GTGAGCCAGA	AATTGATTAA	GGCTTTCTTT	GGTAGGTAAA	2160
TCTAGAGTTT	ATACAGTGTA	CATGTACATA	GTAAAGTATT	TTTGATTAA	AATGTATTTT	2220
AATAACATAT	CTAAAGTCAT	CATGAAGTGG	CTTGATACAT	TTTAAATTCT	TACTCTGGAG	2280
CAACCTACTG	TCTAAGCAGT	TTTGTAATAT	TACTGGTAAT	TGTACAATAC	TTGCATTCCA	2340
GAGTTAAAT	GTTTACTGTA	AATTTTTGTT	CTTTTAAAGA	CTACCTGGGA	CCTGATTTAT	2400
TGAAATTTT	CTCTTTAAAA	ACATTTTCTC	TCGTTAATTT	TCCTTTGTCA	TTTCCTTTGT	2460
TGTCTACATT	AAATCACTTG	AATCCATTGA	AAGTGCTTCA	AGGGTAATCT	TGGGTTTCTA	2520
GCACCTTATC	TATGATGTTT	CTTTTGCAAT	TGGAATAATC	ACTTGGTCAC	CTTGCCCCAA	2580
GCTTTCCCT	CTGAATAAAT	ACCCATTGAA	CTCTGAAAAA	AAAAAAAAAA	AAAA	

Seq ID NO: 294 Protein sequence:  
Protein Accession #: NP\_001489

1	11	21	31	41	51	
MGLLSQGSPL	SWEETKRHAD	HVRRHGILQF	LHIYHAVKDR	HKDVLKWGDE	VEYMLVSFDH	60
ENKKVRLVLS	GEKVLETLQE	KGERTNPNHP	TLWRPEYGSY	MIEGTPGQPY	GGTMSEFNTV	120
EANMRKRRKE	ATSILEENQA	LCTITSFPERL	GCPGFTLPEV	KPNPVEGGAS	KSLFFPDEAI	180
NKHPRFSTLT	RNIRHRRGEK	VVINVPFKD	KNTSPSFIET	FTEDDEASRA	SKPDHIYMDA	240
MGFGMGNCCL	QVTFQACSI	EARYLYDQLA	TICPIVMALS	AASPFYRGYV	SDIDCRWGI	300
SASVDDRTRE	ERGLEPLKNN	NYRISKSRD	SIDSYLSKCG	EKYNDIDLT	DKEIYEQLLQ	360
EGIDHLLAQH	VAHLFIRDPL	TLFEEKIHL	DANESDHFN	IQSTNQWTMR	FKPPPPNSDI	420
GWRVEFRPME	VQLTDFENSA	YVVFVLLTR	VILSYKLDFL	IPLSKVDENM	KVAQKRDAVL	480
QGMFYFRKDI	CKGGNAVVDG	CGKAQNSTEL	AAEYTLMSI	DTIINGKEGV	FPGLIPILNS	540
YLENMEVDVD	TRCSILNYLK	LIKKRASGEL	MTVARWMREF	IANHPDYKQD	SVITDEMNY	600
LILKCNQIAN	ELCECEPELLG	SAFRKVKYSG	SKTDSSN			



Seq ID NO: 295 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 247-816

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
      GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120
10     CCTAGGGGGC ACATTTCCCA CAACTCCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
      GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240
      GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
      CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
      AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
      ATGACAGGAC ATGCTATTCC ACCCAGCCAA TTGGATTCTC AGATTGATGA CTTCACTGGT 480
15     TTCAGCAAAG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
      ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAGC 600
      CAACGAGAAA TTAATGCTGA TATAAACGTT AAATTAGTGA AGGAACTCCG ATGCGTTGGA 660
      CAAAAATATG AAAAAATCTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
      AAGCGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
20     AAGCACCTTA AGAAGAAACT GAAACGTATG ATTTGAGAAT ACTTGTCCCT GGAGGATTAT 840
      CACACCCCAA ATGCATAATC TCGTTAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
      TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCCTGGCAT ATGTTACCGA ATCAAATAGC 960
      CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTC TTTTTCCTCA AGCATTTTAT 1020
      TTGAAAGGAT AACTTGTGTT TTGGTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
25     TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A

```

Seq ID NO: 296 Protein sequence:

Protein Accession #: Eos sequence

```

30     1      11      21      31      41      51
      |      |      |      |      |      |
      MTDKTEKVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
      GHAIPPSQLD SQIDDFTGFS KDRMMQKPGS NAPVGGNVTS SFGDDLECR ETASSPKSQR 120
35     EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR FFESIIEKAA RCMRRDFVKH 180
      LKKKLKRLMI

```

Seq ID NO: 297 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 247-815

```

40     1      11      21      31      41      51
      |      |      |      |      |      |
      AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
      GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120
45     CCTAGGGGGC ACATTTCCCA CAACTCCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
      GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240
      GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
      CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
      AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
      ATGACAGGAC ATGCTATTCC ACCCAGCCAA TTGGATTCTC AGATTGATGA CTTCACTGGT 480
50     TTCAGCAAAG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
      ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAGC 600
      CAACAAGAAA TTAATGCTGA TATAAACGTT AAATTAGTGA AGGAACTCCG ATGCGTTGGA 660
      CAAAAATATG AAAAAATCTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
      AAACGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
      AAGCACCTTA AGAAGAAACT GAAACGTATG ATTTGAGAAT ACTTGTCCCT GGAGGATTAT 840
      CACACCCCAA ATGCATAATC TCATTAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
      TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCCTGGCAT ATGTTACCGA ATCAACTGGC 960
60     CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTC TTTTTCCTCA AGCGTTTTAT 1020
      TTGAAAGGAT AACTTGTGTT TTGGTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
      TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A

```

Seq ID NO: 298 Protein sequence:

Protein Accession #: Eos sequence

```

65     1      11      21      31      41      51
      |      |      |      |      |      |
      MTDKTEKVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
      GHAIPPSQLD SQIDDFTGFS KDRMMQKPGS NAPVGGNVTS SFGDDLECR ETASSPKSQQ 120
70     EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR FFESIIEKAA RCMRRDFVKH 180
      LKKKLKRLMI

```

Seq ID NO: 299 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 247-815

```

75     1      11      21      31      41      51
      |      |      |      |      |      |
      AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
      GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120
80     CCTAGGGGGC ACATTTCCCA CAACTCCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
      GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240
      GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
      CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
      AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420
85     TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480

```



WO 02/086443

PCT/US02/12476

TCAGCAAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540  
 CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAAGCC 600  
 AACAGAAAT TAATGCTGAT ATAAAACGTA AATTAGTGAA GGAACCTCGA TCGGTTGGAC 660  
 5 AAAAAATATGA AAAAAATCTTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720  
 AACGATTTTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTTA 780  
 AGCACCTTAA GAAGAAACTG AAACGTATGA TTTGAGAATA CTTGTCCCTG GAGGATTATC 840  
 ACACCCCAAA TGCATAATCT CATTAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTTT 900  
 TCTACAATGG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTTACCGAA TCAACTGGCC 960  
 10 TTCCAGAGGC TAAGAAATTT CTGTTAGTAA AAGATGTTCT TTTTCCCAA GCGTTTTATT 1020  
 TGAAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCCAC CTGTGCTGGT AGATATTATT 1080  
 AACCATTAG GTAAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 300 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | | |  
 MTDKTEKVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKAKKLMT 60  
 20 GHAIPPSQLD SQIDDFTFGS KDRMMQKPGS NAPVGGNVTS SFGDDLECR ETASSPKSQQ 120  
 EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR FFESIIEKAA RCMRRDFVKH 180  
 LKKKLKRLMI

Seq ID NO: 301 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 247-812

1 11 21 31 41 51  
 | | | | | |  
 30 AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60  
 GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120  
 CCTAGGGGGC ACATTTCCCA CAACTCCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180  
 GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240  
 GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300  
 35 CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360  
 AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAGCTTA 420  
 TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480  
 TCAGCAAAGA TGGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAATGTTA 540  
 CCAGCAATTT CTCTGGAGAT GACCTAGAAT GCAGAGGAAT AGCCTCCTCT CCCAAAAGCC 600  
 40 AACAGAAAT TAATGCTGAT ATAAAATGTC AAGTAGTGAA GGAAATCCGA TGCCTTGGAC 660  
 AATATGAAAA AATCTTCGAA ATGCTTGAAG GAGTGCAAGG ACCTACTGCA GTCAGGAAAC 720  
 GATTTTTTGA ATCCATCATC AAGGAAGCAG CAAGATGTAT GAGACGAGAC TTTGTTAAGC 780  
 ACCTTAAGAA GAAACTGAAA CGTATGATTT GAGAATACTT GTCCCTGGAG GATTATCACA 840  
 CCCCAATGC ATAATCTCAT TAATGATTGA GGAGAGAAAA GGATCAGATT GCTGTTTTCT 900  
 45 ACAATGGAGC AGGATATTGC TGAAGTCTCC TGGCATATGT TACCGAATCA ACTGGCCTTC 960  
 CAGAGGCTAA GAAATTTCTG TTAGTAAAAG ATGTTCTTTT TCCCAAAGCG TTTTATTTGA 1020  
 AAGGATAACT TGTGTTTTGG TTATTTTGTA TTCCACCTG TGCTGGTAGA TATTATTAAC 1080  
 CCATTAGGTA AATACTATTA CAGTCGTGGT TTCTGCA

Seq ID NO: 302 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | | |  
 55 MTDKTEKVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKEKKLMT 60  
 GHAIPPSQLD SQIDDFTFGS KDGMMQKPGS NAPVGGNVTS NFSGDDLECR GIASSPKSQQ 120  
 EINADIKQV VKEIRCLGQY EKIFEMLEGV QGPTAVRKR FESIIEKAAR CMRRDFVKHL 180  
 KKKLKRLMI

Seq ID NO: 303 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 247-815

1 11 21 31 41 51  
 | | | | | |  
 65 AGTGTTCGGC TGGGACAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTTCC ATCCCCCTTG 60  
 GGCCAAACAG GATCGGTGCT TCTGGTGAGA CGTCTCCCCA TGCACATCAC TCCCAGATGC 120  
 CCTAGGGGGC ACATTTCCCA CAACTCCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180  
 GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CCGTCTCTCC TCCAGCAAGG 240  
 70 GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300  
 CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360  
 AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420  
 TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480  
 TCAGCAAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540  
 75 CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAAGCC 600  
 AACAGAAAT TAATGCTGAT ATAAAACCTG AATTAGTGAA GGAACCTCGA TCGGTTGGAC 660  
 AAAAAATATGA AAAAACTTTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720  
 AACGATTTTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTTA 780  
 AGCACCTTAA GAAGAAACTG AAACGTATGA TTTGAGAATA CTTGTCCCTG GAGGATTATC 840  
 80 ACACCCCAAA TGCATAATCT CGTTAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTTT 900  
 TCTACAATGG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTTACCGAA TCAACTGGCC 960  
 TTCCAGAGGC TAAGAAATTT CTGTTAGTAA AAGATGTTCT TTTTCCCAA GCGTTTTATT 1020  
 TGAAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCCAC CTGTGCTGGT AGATATTATT 1080  
 AACCATTAG GTAAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 304 Protein sequence:  
 Protein Accession #: Eos sequence



1	11	21	31	41	51	
MTDKTEKVAV	DPETVFKRPR	ECDSPSYQKR	QRMALLARKQ	GAGDSLIIAGS	AMSKAKKLMT	60
GHAIPPSQLD	SQIDDFTGFS	KDRMMQKPGS	NAPVGGNVTS	SFSGDDLECR	ETASSPKSQQ	120
EINADIKRKL	VKELRCVGQK	YEKIFEMLEG	VQGPTAVRKR	FFESIIEKAA	RCMRRDFVKH	180
LKKKLKRMI						

Seq ID NO: 305 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 87-689

1	11	21	31	41	51	
CGTGGAGGCA	GCTAGCGCGA	GGCTGGGGAG	CGCTGAGCCG	CGCGTCGTGC	CCTGCGCTGC	60
CCAGACTAGC	GAACAATACA	GTCAGGATGG	CTAAAGGTGA	CCCCAAGAAA	CCAAAGGGCA	120
AGATGTCCGC	TTATGCCTTC	TTTGTGCAGA	CATGCAGAGA	AGAACATAAG	AAGAAAAACC	180
CAGAGGTCCC	TGTCAATTTT	GCGGAATTTT	CCAAGAAGTG	CTCTGAGAGG	TGGAAGACGA	240
TGTCCGGGAA	AGAGAAATCT	AAATTTGATG	AAATGGCAAA	GGCAGATAAA	GTGCGCTATG	300
ATCGGGAAAT	GAAGGATTAT	GGACCAGCTA	AGGGAGGCAA	GAAGAAGAAG	GATCCTAATG	360
CTCCCAAAAG	GCCACCGTCT	GGATTCTTCC	TGTTCTGTTC	AGAATTCCGC	CCCAAGATCA	420
AATCCACAAA	CCCCGGCATC	TCTATTGGAG	ACGTGGCAAA	AAAGCTGGGT	GAGATGTGGA	480
ATAATTTAAA	TGACAGTGAA	AAGCAGCCTT	ACATCACTAA	GGCGGCAAG	CTGAAGGAGA	540
AGTATGAGAA	GGATGTTGCT	GACTATAAGT	CGAAAGGAAA	GTTTGATGGT	GCAAAGGGTC	600
CTGCTAAAGT	TGCCCCGAAA	AAGGTGGAAG	AGGAAGATGA	AGAAGAGGAG	GAGGAAGAAG	660
AGGAGGAGGA	GGAGGAGGAG	GATGAATAAA	GAAACTGTTT	ATCTGTCTCC	TTGTGAATAC	720
TTAGAGTAGG	GGAGCGCCGT	AATTGACACA	TCTCTTATTT	GAGAAGTGTC	TGTTGCCCTC	780
ATTAGGTTTA	ATTACAAAAT	TTGATCAGCA	TCATATTGTA	GTCTCTCAAA	GTGCTCTAGA	840
AATTGTCAGT	GGTTTACATG	AAGTGGCCAT	GGGTGTCTGG	AGCACCTGTA	AACTGTATCA	900
AAGTTGTACA	TATTTCCAAA	CATTTTTTAA	ATGAAAAGGC	ACTCTCGTGT	TCTCCTCACT	960
CTGTGCACTT	TGCTGTTGGT	GTGACAAGGC	ATTTAAAGAT	GTTTCTGGCA	TTTTCTTTTT	1020
ATTTGTAAGG	TGGTGGTAAC	TATGGTTATT	GGCTAGAAAT	CCTGAGTTTT	CAACTGTATA	1080
TATCTATAGT	TTGTAAAAAG	AACAAAACAA	CCGAGACAAA	CCCTTGATGC	TCCTTGCTCG	1140
GCGTTGAGGC	TGTGGGGAAG	ATGCCTTTTG	GGAGAGGCTG	TAGCTCAGGG	CGTGCCTGT	1200
GAGGCTGGAC	CTGTTGACTC	TGCAGGGGGC	ATCCATTTAG	CTTCAGGTTG	TCTTGTCTCT	1260
GTATATAGTG	ACATAGCATT	CTGCTGCCAT	CTTAGCTGTG	GACAAAGGGG	GGTCAGCTGG	1320
CATGAGAATA	TTTTTTTTTT	TAAGTGCAGT	AGTTTTTAAA	CTGTTTGTTC	TTAAACAAAC	1380
TATAGAACTC	TTCATTGTCA	GCAAAGCAAA	GAGTCACTGC	ATCAATGAAA	GTTCAAGAAC	1440
CTCCTGTACT	TAAACACGAT	TCGCAACGTT	CTGTTATTTT	TTTGTATGTC	TTAGAATGCT	1500
GAAATGTTTT	TGAAGTTAAA	TAAACAGTAT	TACATTTTTA	AAACTCTTCT	CTATTATAAC	1560
AGTCATAATTC	TGACTCACAG	CAGTGAACAA	ACCCCACTC	CATTGTATTT	GGAGACTGGC	1620
CTCCCTATAA	ATGTGGTAGC	TTCTTTTTATT	ACTCAGTGGC	CAGCTCACTT	AGGGCTGAGA	1680
TGAAGGAGAG	GGCTACTTGA	AGCTACTGTG	TGATTTTGTG	TGTGTCTGAG	TGGCATTGAG	1740
ATGAAGTCTG	GAGGAGTTAG	GAGAACGACA	TAGGCAAGGT	TCAGCAGCCT	TCCAAGGTAT	1800
AGGAAGGTGG	GTGATTAGGA	CTGAGGCTAT	CTAGGTTTAA	CTTTGTGCCC	ACCTCCACCC	1860
CCTATTTTGT	GGGGCCAAAT	GCATTGCTAA	ACAGCAATTT	CAGAGTGTAT	GGTGTGTCAA	1920
AAATTAAGGC	CTTATTGTTT	TTCTCTTTCA	CCCCACCC	CCGTGCTCCT	GGCACATATC	1980
ACATTATTTG	TGGTGCCCAA	CATTGGCCGT	CTTGAGCCTG	CTGCTGGTCT	CCTGGATGCC	2040
AGTGAGGGTA	TGTGGGATGG	GGTGGTGGGG	TAGGGGACGG	TATCCTTTTT	TTGCTCCTAC	2100
TTGGAAACAC	CAAACACCCC	AAGGAAGATG	ATAGGCTCCA	TCTTGGGCCA	CCTGAGCTAT	2160
AGGGCAGGCT	AATGGAATCA	ACCATTTCTG	AGCACTAAAT	GTATCATGAA	AAGTTGAATG	2220
GCCTGCTCAT	AAGTTTAGCT	CATTCACTGG	AAATGTAGAT	TGATGTTCAA	TGTTAAACTG	2280
GAAGGAGCTT	GGTTTGTTGT	TCAGTGGTTA	TATTAGTGGG	TAGTGTAAAC	TTTTATCCAG	2340
GTGGGGGTGA	GGGGAGATGG	CCACAGTAGC	AAGTGGTGAC	ACTAAATACC	ATTTTGAAGG	2400
CTGATGTGTA	TATACATCAT	TACTGTCCGT	AGCAATGAAG	GATACAGTAC	TGTGTTGTGG	2460
GTGAGTGTGG	CTATTGCCCA	GCATTAATAT	TTGGGTGTGT	ATGTTTGAGG	CTATGAAACA	2520
CGCAGGAGTG	TTTTTGTGCT	ATTAATTTTA	AGAGAAAGCA	GCTTTTCTCT	AAAATTCACT	2580
GTTGAGAAAC	TTGCATGTCT	GGAGGCGGTG	TCCTCTCCGC	CCTGTCGGGT	CCTGGATGAG	2640
TACGAGTTAT	GGTCACGGTC	ACAGCCTGAT	CTCTTATGTG	TTCATAGCCA	TTGCTCTCTC	2700
CATCAGAACT	GTTTGTCTCT	AATGTGTTCC	TCTAGTTCTA	GAAAATGACC	ACTAATTTAA	2760
AAAACCTCGT	TGTGAGGTTT	GCCCAGAGGC	ACTTGTTCCT	GAATTTCCCC	TCCTGCTTCA	2820
GCCATGTCCT	TGTCACCTGG	CATTCTAAGC	TAAAGCTTTA	GCTTCCCAAT	TCGTGATGTG	2880
CTAGGCCAAG	ATTCCGGGAGC	TGTTGCCAGC	CTCGTCAAAT	ATGGAAGAGA	AACAACCTGC	2940
GGTCAAAAGG	GAGTGATTGG	TTAAGTGGTG	CGCGTCTATC	TCATAACTAG	ATGTACCAAC	3000
CAGGGAAGGG	CCAAGGATGG	AAAGGGGTAA	CTTTTGTGCT	TCCAAAGTAG	CTAAGCAGAA	3060
GTGGGGGAGC	AGTTTAGCCA	GATGATCTTT	GATTAGGCAA	ACATTGAGTT	TTAAAGAGGC	3120
TGTCAAGTTG	AGGCCACTTG	GTCCATTAGC	TGGGGCAGCA	AGATCACTAC	TCAACGTTTT	3180
CACACTGTGG	CAAGATTGCT	CTTCTAGTGG	AATAATGCCC	TAGTTTCTCT	GAGATGATGT	3240
AAGTGGCATG	ATGTTACCTA	AGGCTTAGGC	TTAGCTTGAT	TTCTGGGCCC	ACTGTCTGTG	3300
TTCTTAAGAT	GCCAACCTGT	TGCTTTTTTT	TTTTTTTTTCC	CCCATTATAA	AGGATAGTAC	3360
CTACTCCCTC	TAACCACCTC	ACCCCATTTCT	TGAATGACAT	TTTATCCTTC	GGAAAGAACA	3420
AGGCTGTGAT	GTAGTGACTA	TTGTCTGTGT	CTCCTGTGTG	TGTCTGTTCT	TGTCACAAAT	3480
GTATTTGGGG	ACGTTGGATG	CATTCAATTT	CTGTAATAAA	G		

Seq ID NO: 306 Protein sequence:  
Protein Accession #: NP\_005333.1

1	11	21	31	41	51	
MAKGDPKKPK	GKMSAYAFFV	QTCREEHKKK	NPEVPVNFAB	FSKKCSERWK	TMSGKEKSKF	60
DEMAKADKVR	YDREMKDYG	AKGGKKKKDP	NAPKRPPSGF	FLFCSEFRPK	IKSTNPGISI	120
GDVAKKLGE	WNNLNDSEKQ	PYITKAALKK	EKYEKDVADY	KSKGKFDGAK	GPAKVARKKV	180
EEEEEEEEEE	EEEEEEEEEE					

Seq ID NO: 307 DNA sequence  
Nucleic Acid Accession #: NM\_022342  
Coding sequence: 1..2178



	1	11	21	31	41	51	
5	ATGGGTACTA	GGAAAAAAGT	TCATGCATTT	GTCCGTGTCA	AACCCACCGA	TGACTTTGCT	60
	CATGAAATGA	TCAGATACGG	AGATGACAAA	AGAAGCATTG	ATATTCACCT	AAAAAAAGAC	120
	ATTCGGAGAG	GAGTTGTCAA	TAACCAACAG	ACAGACTGGT	CGTTTAAGTT	GGATGGAGTT	180
	TTCACGATG	CCTCCCAGGA	CTTGGTTTAT	GAGACAGTTG	CAAAGGATGT	GGTTTCTCAG	240
	CCCTCGATG	GCTATAATGG	CACCATCATG	TGTTATGGGC	AGACGGGAGC	TGGCAAGACA	300
10	ACACCATGA	TGGGGGCAAC	TGAGAATTAC	AAGCACCAGG	GGATCCTCCC	TCGTGCCCTG	360
	AGCAGGTTT	TTAGGATGAT	CGAAGAACGC	CCCACACATG	CCATCACTGT	GCGTGTTCCT	420
	ACTTGGAAA	TCTATAATGA	GAGCCTGTTT	GATCTCCTGT	CCACTCTGCC	CTATGTTGGA	480
	CCTCAGTCA	CACCAATGAC	CATCGTGGAA	AACCCTCAAG	GAGTCTTCAT	TAAGGGCTTG	540
	CAGTTCACC	TCACAAGTCA	GGAGGAGGAT	GCATTACAGC	TCCTTTTGA	GGGTGAGACC	600
15	ACAGGATTA	TAGCCTCCCA	CATATGAAC	AAAACTCTT	CCAGATCACA	CTGCATTTTC	660
	CCATCTACT	TAGAGGCCCA	TTCCCGGACC	TTATCAGAGG	AAAAGTACAT	CACTTCCAAA	720
	TTAACTTGG	TGGATCTGGC	AGGCTCAGAG	AGGCTGGGGA	AGTCTGGGTC	TGAGGGCCAA	780
	TCCTGAAGG	AAGCCACCTA	CATCAACAAA	TCGCTCTCAT	TCCTGGAGCA	GGCCATCATT	840
	CCCTTGGGG	ACCAGAAGCG	GGACCACATC	CCCTTTCGGC	AGTGCAAGCT	CACCCACGCT	900
20	TGAAGGACT	CGTTAGGGGG	AAACTGCAAT	ATGGTCTCG	TGACAAACAT	CTATGGAGAA	960
	CTGCCCAGT	TAGAAGAAAC	GCTATCTTCA	CTGAGATTTG	CCAGCAGGAT	GAAGCTAGTC	1020
	CCACTGAGC	CTGCCATCAA	TGAAAAGTAT	GATGCTGAGA	GAATGGTCAA	GAACCTGGAG	1080
	AGGAAGTAG	CACACTCAA	GCAGGAGCTG	GCTATCCATG	ACAGCCTGAC	CAACCGCACC	1140
	TTGTGACCT	ATGACCCCAT	GGATGAAATC	CAGATTGCTG	AGATCAACTC	CCAGGTGCGG	1200
25	GGTACCTGG	AGGGGACACT	GGACGAGATC	GACATAATCA	GCCTTAGACA	GATCAAGGAG	1260
	TGTTCAACC	AGTTCCGGGT	GGTTCTGAGC	CAACAGGAAC	AGGAAGTGGA	GTCCACTTTG	1320
	GCAGGAAGT	ACACCCTCAT	TGACAGGAAT	GACTTTGCAG	CCATTTCTGC	TATCCAGAAG	1380
	CGGGGCTTG	TGGATGTTGA	TGGCCACCTA	GTGGGTGAGC	CTGAAGGACA	AAACTTTGGA	1440
	TCGGAGTCG	CCCCTTTCTC	TACCAACCTT	GGGAAGAAAG	CCAAGTCCAA	GAAGACATTC	1500
30	AAGAGCCAC	TCAGGCCCGA	CACCCACCCC	TCCAAACCAG	TGGCCTTTGA	GGAGTTTAAG	1560
	ATGAGCAAG	GTAGTGAGAT	CAACCGAATT	TTCAAAGAAA	ACAAATCCAT	CTTGAATGAA	1620
	GGAGGAAAA	GGGCCAGCGA	GACCACACAG	CACATCAATG	CCATCAAGCG	GGAGATTGAT	1680
	TGACCAAGG	AGGCCCTGAA	TTTCCAGAAG	TCACTACGGG	AGAAGCAAGG	CAAGTACGAA	1740
	ACAAGGGGC	TGATGATCAT	CGATGAGGAA	GAATTCCTGC	TGATCCTCAA	GCTCAAAGAC	1800
35	TCAAGAAGC	AGTACCGCAG	CGAGTACCAG	GACCTGCGTG	ACCTCAGGGC	TGAGATCCAG	1860
	ATTGCCAGC	ACCTAGTGGA	TCAGTGTGCG	CACCGCCTGC	TCATGGAATT	TGACATCTGG	1920
	ACAATGAGT	CCTTTGTCTAT	CCCTGAGGAC	ATGCAGATGG	CACTGAAGCC	AGGCGGCAGC	1980
	TCCGGCCAG	GCATGGTCCC	TGTGAACAGG	ATTGTGTCTC	TGGGAGAAGA	TGACCAGGAC	2040
	AATTTCAGCC	AGCTGCAGCA	GAGGGTGCTT	CCTGAGGGCC	CTGATTCCAT	CTCCTTCTAC	2100
40	ATGCCAAAG	TCAAGATAGA	GCAGAAGCAT	AATTACTTGA	AAACCATGAT	GGGCCTCCAG	2160
	AGGCACATA	GAAAATAG					

Seq ID NO: 308 Protein sequence:

Protein Accession #: NP\_071737

	1	11	21	31	41	51	
45	MGTRKKVHAF	VRVKPTDDFA	HEMIRYGDDK	RSIDIHLKKD	IRRGVVMNQ	TDWSFKLDGV	60
	LHDASQDLVY	ETVAKDVVSQ	ALDGYNGTIM	CYQGTGAGKT	YTMMGATENY	KHRGILPRAL	120
	QQVFRMIEER	PTHAITVRVS	YLEIYNESLF	DLLSTLPYVG	PSVTPMTIVE	NPQGVFIKGL	180
50	SVHLTSQEED	AFSLLEFGET	NRIIASHTMN	KNSSRSHCIF	TIYLEAHSRT	LSEKEYITSK	240
	INLVDLAGSE	RLGKSGSEGO	VLKEATYINK	SLSFLEQAI	ALGDQKRDHI	PFRQCKLTHA	300
	LKDSLGGNCN	MVLVTNIYGE	AAQLEETLSS	LRFASRMKLV	TTEPAINEKY	DAERMVKNLE	360
	KELALLKQEL	AIHDSLNTRT	FVTYDPMDEI	QIAEINSQVR	RYLEGTLDDEI	DIISLRQIKE	420
55	VFNQFRVLS	QQEQEVESL	RRKYTLIDRN	DFAAISAIQK	AGLVDVDGHL	VGEPEGQNF	480
	LGVAPFSTKP	GKKAKSKKTF	KEPLRPDTPP	SKPVAFEFEK	NEQGSEINRI	FKENKSILNE	540
	RRKRASETQ	HINAIKREID	VTKEALNFQK	SLREKQKYE	NKGLMIIDEE	EFLILKLKD	600
	LKKQYRSEYQ	DLRDLRAEIQ	YCQHLVDQCR	HRLLMEFDIW	YNESFVIPED	MQMALKPFGS	660
	IRPGMVPVNR	IVSLGEDDQD	KFSQLQQRVL	PEGPDSISFY	NAVKIEQKH	NYLKTMMGLQ	720
60	QAHRK						

Seq ID NO: 309 DNA sequence

Nucleic Acid Accession #: CAT cluster

	1	11	21	31	41	51	
65	TTTTTTTTTT	TTTTTTTTTA	TGCCTGCTGT	CATGCTCTGT	CTACCAGGGT	GAATTTCCAA	60
	AAATTTCTGC	ATAGCAATTT	TAGCCAAAAC	TATATATGTT	CTGGGGAGGA	TAGGCATAGG	120
	CACATTGAAG	ACCAAAGGAA	AGAGTGAAGA	AGTGTAGTTG	GGTCATTGTG	AATGGATGTT	180
70	TAGATTGTCA	AGAAAAGTGG	GCCAGAGGCC	CCACCTCACA	CTAGGACGGC	AATTGCCTCT	240
	CATTAGTATC	TCAGGCACCA	TGGGTCTTAT	TTGGTGTCTAT	AAGAAACACC	CTCAACAAAG	300
	TAATGAACCC	TCAGCCTCCA	GCTTCTCTTC	TTCGGGATTG	TTCTTAGGGC	CTCCTTTTTC	360
	CTTTTATGTT	TCCAGTACCC	TGAATTTCTT	ATTCCCATCC	CCCATTAATA	TCTGCTTCAA	420
	AGAAAAACA	AGAAGGACAC	ATTCACTTTA	AGATCCAAAT	GAATGATAAG	AGCTTAAAC	480
75	ATTATACTTA	TCAGTATTAT	TTGCATTTTT	ATAGAAACCA	AAACCATATT	TCAACAAC	

Seq ID NO: 310 DNA sequence

Nucleic Acid Accession #: NM\_018622.2

Coding sequence: 1-1140

	1	11	21	31	41	51	
80	ATGGCGTGGC	GAGGCTGGGC	GCAGAGAGGC	TGGGGCTGCG	GCCAGGCGTG	GGGTGCGTGC	60
	GTGGGCGGGC	GCAGCTGCGA	GGAGCTCACT	GCGGCTCTAA	CCCCGCCGCA	GCTCCTCGGA	120
	CGAGGTTTA	ACTTCTTTAT	TCAACAAAAA	TGCGGATTCA	GAAAAGCACC	CAGGAAGGTT	180
85	GAACCTCGAA	GATCAGACCC	AGGGACAAGT	GGTGAAGCAT	ACAAGAGAAG	TGCTTTGATT	240
	CCTCCTGTGG	AAGAAACAGT	CTTTTATCCT	TCTCCCTATC	CTATAAGGAG	TCTCATAAAA	300
	CCTTTATTTT	TTACTGTTGG	GTTTACAGGC	TGTGCATTTG	GATCAGCTGC	TATTTGGCAA	360



TATGAATCAC TGAAATCCAG GGTCCAGAGT TATTTTGTATG GTATAAAAGC TGATTGGTTG 420  
GATAGCATAA GACCACAAA AGAAGGAGAC TTCAGAAAAGG AGATTAACAA GTGGTGGAAT 480  
AACCTAAGTG ATGGCCAGCG GACTGTGACA GGTATTATAG CTGCAATGT CCTTGTATTC 540  
TGTTTATGGA GAGTACCTTC TCTGCAGCGG ACAATGATCA GATATTTTAC ATCGAATCCA 600  
GCCTCAAAGG TCCTTTGTTC TCCAATGTTG CTGTCAACAT TCAGTCACTT CTCCTTATTT 660  
CACATGGCAG CAAATATGTA TGTTTTGTGG AGCTTCTCTT CCAGCATAGT GAACATTCTG 720  
GGTCAAGAGC AGTTCATGGC AGTGTACCTA TCTGCAGGTG TTATTTCCAA TTTTGTTCAGT 780  
TACCTGGGTA AAGTTGCCAC AGGAAGATAT GGACCATCAC TTGGTGCAATC TGGTGCCATC 840  
ATGACAGTCC TCGCAGCTGT CTGCACTAAG ATCCCAGAAG GGAGGCTTGC CATTATTTTC 900  
CTTCCGATGT TCACGTTTAC AGCAGGGAAT GCCCTGAAAG CCATTATCGC CATGGATACA 960  
GCAGGAATGA TCCTGGGATG GAAATTTTTT GATCATGCGG CACATCTTGG GGGAGCTCTT 1020  
TTTGGAATAT GGTATGTTAC TTACGGTCAT GAACTGATT GGAAGAACAG GGAGCCGCTA 1080  
GTGAAAATCT GGCATGAAAT AAGGACTAAT GGCCCCAAAA AAGGAGGTGG CTCTAAGTAA

Seq ID NO: 311 Protein sequence:  
Protein Accession #: NP\_061092.2

1 11 21 31 41 51  
MAWRGWAQRG WCGQAWGAS VGGRSCEELT AVLTPPQLLG RRFNFFIQQK CGFRKAPRKV 60  
EPRRSDPGTS GEAYKRSALI PVVEETVFYP SPYPIRSLIK PLFFTVGFTG CAFGSAAIWQ 120  
YESLKSQRVS YFDGIKADWL DSIRPQKEGD FRKEINKWVN NLSGQRTVT GIIAANVLVF 180  
CLWRVPSLQR TMIRYFTSNP ASKVLCSFML LSTFSHFSLF HMAANMYVLW SFSSSIVNIL 240  
GQEQFMAYVL SAGVISNFVS YLGKVTATRY GPSLGASGAI MTLAAVCTK IPEGRLAIF 300  
LPMFTFTAGN ALKAIAMDT AGMILGWKFF DHAAHLGGAL FGIWYVTYGH ELIWKNNREPL 360  
VKIWEHIRTN GPKKGGGSK

Seq ID NO: 312 DNA sequence  
Nucleic Acid Accession #: NM\_000625  
Coding sequence: 195..3656

1 11 21 31 41 51  
CTCTCGGCCA CCTTTGATGA GGGGACTGGG CAGTTCTAGA CAGTCCCGAA GTTCTCAAGG 60  
CACAGGTCTC TTCCTGGTTT GACTGTCTTT ACCCGGGGGA GGCAGTGCAAG CCAGCTGCAA 120  
GCCCCACAGT GAAGAACATC TGAGCTCAAA TCCAGATAAG TGACATAAGT GACCTGCTTT 180  
GTAAAGCCAT AGAGATGGCC TGTCTTGGGA AATTTCTGTT CAAGACCAA TTCCACCAGT 240  
ATGCAATGAA TGGGGAAAAA GGCATCAACA ACAATGTGGA GAAAGCCCCC TGTGCCACCT 300  
CCAGTCCAGT GACACAGGAT GACCTTCAGT ATCACAACCT CAGCAAGCAG CAGAATGAGT 360  
CCCCGCAGCC CCTCGTGGAG ACGGGAAAGA AGTCTCCAGA ATCTCTGGTC AAGCTGGATG 420  
CAACCCCATG GTCCTCCCA CGGCATGTGA GGATCAAAAA CTGGGCGAGC GGGATGACTT 480  
TCCAAGACAC ACTTCACCAT AAGGCCAAAG GGATTTTAAC TTGCAAGTCC AAATCTTGCC 540  
TGGGGTCCAT TATGACTCCC AAAAGTTTGA CCAGAGGACC CAGGGACAAG CCTACCCCTC 600  
CAGATGAGCT TCTACCTCAA GCTATGTAAC TTGTCAACCA ATATTACGGC TCCCTCAAAG 660  
AGGCAAAAT AGAGGAACAT CTGGCCAGGG TGAAGCGGT AACAAAGGAG ATAGAAACAA 720  
CAGTAACCTA CCAACTGACG GGAGATGAGC TCATCTTCGC CACCAAGCAG GCCTGGCGCA 780  
ATGCCCCACG CTGCATTGGG AGGATCCAGT GGTCCAACCT GCAGGTCTTC GATGCCCGCA 840  
GCTGTTCCAC TGCCCGGGA ATGTTTGAAC ACATCTGCAG ACACGTGCGT TACTCCACCA 900  
ACAATGGCAA CATCAGGTCC GCCATCACCC TGTTCCCCCA GCGGAGTGAT GGCAAGCACG 960  
ACTTCCGGGT GTGGAATGCT CAGCTCATCC GCTATGCTGG CTACCAGATG CCAGATGGCA 1020  
GCATCAGAGG GGACCTGCC AACGTGGAAT TCACTCAGCT GTGCATCGAC CTGGGCTGGA 1080  
AGCCCAAGTA CGGCCGCTTC GATGTGGTCC CCCTGGTCTC GCAGGCCAAT GGCCGTGACC 1140  
CTGAGCTCTT CGAAATCCCA CCTGACCTTG TGCTTGAGGT GGCCATGGAA CATCCCAAAT 1200  
ACGAGTGGTT TCGGGAACGT GAGCTAAAGT GGTACGCCCT GCCTGCAGTG GCCAACATGC 1260  
TGCTTGAGGT GGGCGGCTTG GAGTTCCAG GGTGCCCTT CAATGGCTGG TACATGGGCA 1320  
CAGAGATCGG AGTCCGGGAC TTCTGTGATG TCCAGCGCTA CAACATCCTG GAGGAAGTGG 1380  
GCAGGAGAAT GGGCTTGGAA ACGCATGAGT TGGCTCTGCT CTGGAAGAC CAGGCTGTCTG 1440  
TTGAGATCAA CATTGCTGTG CTCCATAGTT TCCAGAAGCA GAATGTGACC ATCATGGACC 1500  
ACCACTCGGC TGCAGAAATC TTCATGAAGT ACATGCAGAA TGAATACCGG TCCCGTGGGG 1560  
GCTGCCCGGC AGACTGGATT TGGCTGGTCC CTCCCATGTC TGGGAGCATC ACCCCCGTGT 1620  
TTCACCAGGA GATGCTGAAC TACGTCTGT CCCCTTTCTA CTACTATCAG GTAGAGGCCT 1680  
GGAAAACCCA TGTCTGGCAG GACGAGAAGT GGAGACCAA GAGAAGAGAG ATTCCATTGA 1740  
AAGTCTTGGT CAAAGCTGTG CTCTTTGCTT GTATGCTGAT GCGCAAGACA ATGGCGTCCC 1800  
GAGTCAGAGT CACCATCCTC TTTGCGACAG AGACAGGAAA ATCAGAGGCG CTGGCCTGGG 1860  
ACCTGGGGGC CTTATTCAGC TGTGCCTTCA ACCCAAGGT TGTCTGCATG GATAAGTACA 1920  
GGCTGAGCTG CCTGGAGGAG GAACGGCTGC TGTGGGTGGT GACCAGTACG TTTGGCAATG 1980  
GAGACTGCCC TGGCAATGGA GAGAAACTGA AGAAATCGCT CTTATGCTG AAAGAGCTCA 2040  
ACAACAAAT CAGGTACGCT GTGTTTGGCC TCGGCTCCAG CATGTACCTT CCGTTCTGCG 2100  
CCTTTGCTCA TGACATTGAT CAGAAGCTGT CCCACCTGGG GGCTCTCAG CTCACCCCGA 2160  
TGGGAGAAGG GGATGAGCTC AGTGGGAGG AGGACGCTT CCGCAGCTGG GCCGTGCAA 2220  
CCTTCAAGGC AGCCTGTGAG ACCTTTGATG TCCGAGGCAA ACAGCACATT CAGATCCCCA 2280  
AGCTCTACAC CTCCAATGTG ACCTGGGACC CGCACCATA CAGGCTCGTG CAGGACTCAC 2340  
AGCCTTTGGA CCTCAGCAA GCCCTCAGCA GCATGCATGC CAAGAACGTG TTCACCATGA 2400  
GGCTCAAATC TCGGCAGAA CTACAAAGTC CGACATCCAG CCGTGCCACC ATCCTGGTGG 2460  
AACTCTCCTG TGAGGATGGC CAAGGCCTG ACTACCTGCC GGGGGAGCAC CTTGGGGTTT 2520  
GCCAGGCAA CCGCCGCGC CTGGTCCAAG GCATCCTGGA GCGAGTGGTG GATGGCCCCA 2580  
CACCCACCA GGCAGTGCCG CTGGAGGCC TGGATGAGAG TGGCAGCTAC TGGGTGAGT 2640  
ACAAGAGGCT GCGCCCTGCT TCACTCAGCC AGGCCCTCAC CTACTTCTG GACATCACC 2700  
CACCCCAAC CCAGCTGCTG CTCCAAAAGC TGGCCAGGT GGCCACAGAA GAGCCTGAGA 2760  
GACAGAGGCT GGAGGCCCTG TGCCAGCCCT CAGAGTACAG CAAGTGGAA TACCAACA 2820  
GCCCCACATT CCTGGAGGTG CTAGAGGAG TCCCCTCCT GCGGGTGTCT GCTGGCTTCC 2880  
TGCTTTCCCA GCTCCCATTT CTGAAGCCCA GGTCTTACTC CATCAGCTCC CCCCAGGATC 2940  
ACACGCCCAC GGAGATCCAC CTGACTGTGG CCGTGGTCAC CTACCACACC CGAGATGGCC 3000  
AGGGTCCCCT GCACCACGGC GTCTGCAGCA CATGGCTCAA CAGCCTGAAG CCCCAGACC 3060  
CAGTGCCCTG CTTTGTGCGG AATGCCAGCG GCTTCCACCT CCCCAGGAT CCCTCCCATC 3120



5 CTTGCATCCT CATCGGGCCT GGCACAGGCA TCGCGCCCTT CCGCAGTTTC TGGCAGCAAC 3180  
 GGCTCCATGA CTCCAGCAC AAGGGAGTGC GGGGAGGCCG CATGACCTTG GTGTTTGGGT 3240  
 GCCGCCGCC AGATGAGGAC CACATCTACC AGGAGGAGAT GCTGGAGATG GCCAGAAGG 3300  
 GGGTGCTGCA TGCCTGTCAC ACAGCCTATT CCCGCTGCCC TGGCAAGCCC AAGGTCTATG 3360  
 TTCAGGACAT CCTGCGGCAG CAGCTGGCCA GCGAGGTGCT CCGTGTGCTC CACAAGGAGC 3420  
 CAGGCCACCT CTATGTTTGC GGGGATGTGC GCATGGCCCG GGACGTGGCC CACACCCTGA 3480  
 AGCAGCTGGT GGCTGCCAAG CTGAAATTGA ATGAGGAGCA GGTGAGGAC TATTTCTTTC 3540  
 AGCTCAAGAG CCAGAAGCGC TATCAGGAAG ATATCTTTGG TGCTGTATTT CCTTACGAGG 3600  
 10 CGAAGAAGGA CAGGGTGGCG GTGCAGCCCA GCAGCCTGGA GATGTCAGCG CTCTGAGGGC 3660  
 CTACAGGAGG GGTAAAGCT GCCGGCACAG AACTTAAGGA TGGAGCCAGC TCTGCATTAT 3720  
 CTGAGGTCAC AGGGCCTGGG GAGATGGAGG AAAGTGATAT CCCCAGCCT CAAGTCTTAT 3780  
 TTCCTCAACG TTGCTCCCA TCAAGCCCTT TACTTGACCT CCTAACAAGT AGCACCTTGG 3840  
 ATTGATCGGA GCCTC

15 Seq ID NO: 313 Protein sequence:  
 Protein Accession #: NP\_000616

20 1 11 21 31 41 51  
 | | | | | |  
 MACPWKFLFK TKFHQYAMNG EKGINNVEK APCATSSPVT QDDLQYHNLS KQONESPQPL 60  
 VETGKKSPES LVKLDATPLS SPRHVRIKNW GSGMTFQDTL HHKAKGILTC RSKSCLGSIM 120  
 TPKSLTRGPR DKPTPPDELL PQAIEFVNQY YGSLKEAKIE EHLARVEAVT KEIETTVTYQ 180  
 25 LTGDELIFAT KQAWRNAPRC IGRIQWSNLQ VFDARSCSTA REMFEHICRH VRYSTNNGNI 240  
 RSAITVFPQR SDGKHDFRVW NAQLIRYAGY QMPDGSIRGD PANVEFTQLC IDLGWPKPYG 300  
 RFDVVPVLVQ ANGRDPELFE IPPDLVLEVA MEHPKYEWFR ELELKWYALP AVANMLLEV 360  
 GLEFPGCPFN GWYMGTEIGV RDFCDVQRYN ILEEVGRRMG LETHKLASLW KDQAVVEINI 420  
 AVLHSFQKQN VTIMDHHSAA ESFMKYMONE YRSRGGCPAD WIWLVPPMSG SITPVFHQEM 480  
 30 LNYVLSPPFY YQVEAWKTHV WQDEKRRPKR REIPLKVLVK AVLFACMLMR KTMSRVRVT 540  
 ILFATETGKS EALAWDLGAL FSCAFNPKVV CMDKYRLSCL EEERLLLVVT STFGNGDCPG 600  
 NGEKLKKSFL MLKELNNKFR YAVFGLGSSM YPRFCAFAHD IDQKLSHLGA SOLTPMGED 660  
 ELSGQEDAFR SWAVQTFKAA CETFDVRGKQ HIQIPKLYTS NVTWDPHHYR LVQDSQPLDL 720  
 SKALSSMHAK NVFTMLKSR QNLQSPTSSR ATILVELSCE DGQGLNYLPG EHLGVCPGNQ 780  
 35 PALVQGILER VVDGTPHQA VRLEALDESG SYWVSDKRLP PCSLSQALTY FLDITTPPTQ 840  
 LLLQKLAQVA TEEPERQRLE ALCQPSEYSK WKFTNSPTFL EVLEEFPSLR VSAGFLLSQL 900  
 PILKPRFYSI SSPRDHTPTE IHLTVAVVTY HTRDGQGPLH HGVCSTWLNS LKPQDPVPCF 960  
 VRNASGFHLP EDPSPHCILI GPGTGIAPFR SFWQQLHDS QHKGVRGGRM TLVFGCRRPD 1020  
 EDHIYQEEML EMAQKGVLHA VHTAYSRLPG KPKVYVQDIL RQQLASEVLR VLHKEPGHLY 1080  
 40 VCGDVRMARD VAHTLKQLVA AKLKLNEEQV EDYFFQLKSQ KRYHEDIFGA VFPYEAKKDR 1140  
 VAVQPSLEM SAL

Seq ID NO: 314 DNA sequence  
 Nucleic Acid Accession #: XM\_087254  
 Coding sequence: 47..2332

45 1 11 21 31 41 51  
 | | | | | |  
 AGAGTACGTG TTTACAGATA AACTGGTAC ACTGACAGAA AATGAGATGC AGTTTCGGGA 60  
 50 ATGTTCAATT AATGGCATGA AATACCAAGA AATTAATGGT AGACTTGATC CCGAAGGACC 120  
 AACACCAGAC TCTTCAGAAAG GAAACTTATC TTATCTTAGT AGTTTATCCC ATCTTAACAA 180  
 CTTATCCCAT CTTACAACCA GTTCCTCTTT CAGAACCAGT CCTGAAAATG AAAGTGAAC 240  
 AATTAAAGAA CATGATCTCT TCTTTAAAGC AGTCAGTCTC TGTCACACTG TACAGATTAG 300  
 55 CAATGTTCAA ACTGACTGCA CTGGTGATGG TCCCTGGCAA TCCAACCTGG CACCATCGCA 360  
 GTTGGAGTAC TATGCATCTT CACCAGATGA AAAGGCTCTA GTAGAAGCTG CTGCAAGGAT 420  
 TGGTATTGTG TTTATTGGCA ATTCTGAAGA AACTATGGAG GTTAAAACTC TTGGAAGAACT 480  
 GGAACGGTAC AACTGCTTC ATATTCTGGA ATTTGATTCA GATCGTAGGA GAATGAGTGT 540  
 AATTGTTTCA GCACCTTCAG GTGAGAAAGT ATTATTTGCT AAAGGAGCTG AGTCATCAAT 600  
 60 TCTCCCTAAA TGTATAGGTG GAGAAATAGA AAAAACAGAG ATTCATGTAG ATGAATTTGC 660  
 TTTGAAAGGG CTAAGAACTC TGTGTATAGC ATATAGAAAA TTTACATCAA AAGAGTATGA 720  
 GGAAATAGAT AAACGCATAT TTGAAGCCAG GACTGCCTTG CAGCAGCGGG AAGAGAAATT 780  
 GGCAGCTGTT TTCCAGTTCA TAGAGAAAGA CCTGATATTA CTTGGAGCCA CAGCAGTAGA 840  
 AGACAGACTA CAAGATAAAG TTCGAGAAAC TATTGAAGCA TTGAGAAATGG CTGGTATCAA 900  
 AGTATGGGTA CTTACTGGGG ATAAACATGA AACAGCTGTT AGTGTGAGTT TATCATGTGG 960  
 65 CCATTTTCAT AGAACCATGA ACATCCTTGA ACTTATAAAC CAGAAATCAG ACAGCGAGTG 1020  
 TGCTGAACAA TTGAGGCAGC TTGCCAGAAG AATTACAGAG GATCATGTGA TTCAGCATGG 1080  
 GCTGGTAGTG GATGGGACCA GCCTATCTCT TGCCTCAGG GAGCATGAAA AACTATTTAT 1140  
 GGAAGTTTGC AGAAATTGTT CAGCTGTATT ATGCTGTCTG ATGGCTCCAC TGCAGAAAGC 1200  
 70 AAAAGTAATA AGACTAATAA AAATATCACC TGAGAAACCT ATAACATTGG CTGTTGGTGA 1260  
 TGGTGCTAAT GACGTAAGCA TGATACAAGA AGCCCATGTT GGCATAGGAA TCATGGGTAA 1320  
 AGAAGGAAGA CAGGCTGCAA GAAACAGTGA CTATGCAATA GCCAGATTTA AGTTCTCTCT 1380  
 CAAATTGCTT TTTGTTTCATG GTCATTTTTA TTATATTAGA ATAGCTACCC TTGTACAGTA 1440  
 TTTTTTTTAT AAGAATGTGT GCTTTATCAC ACCCCAGTTT TTATATCAGT TCTACTGTTT 1500  
 75 GTTTTCTCAG CAAACATTGT ATGACAGCGT GTACCTGACT TTATACAATA TTTGTTTTTAC 1560  
 TTCCCTACCT ATTCTGATAT ATAGTCTTTT GGAACAGCAT GTAGACCTCT ATGTGTTTACA 1620  
 AAATAAGCCC ACCCTTTATC GAGACATTAG TAAAAACCGC CTCCTAAGTA TTTAAACATT 1680  
 TCTTTATTGG ACCATCCTGG GCTTCAGTCA TGCCTTTATT TTCTTTTGTG GATCCTATTT 1740  
 ACTAATAGGG AAAGATACAT CTCTGCTTGG AAATGGCCAG ATGTTTGGAA ACTGGACATT 1800  
 80 TGGCACTTTG GTCTTCACAG TCATGGTTAT TACAGTCACA GTAAAGATGG CTCTGGAAC 1860  
 TCATTTTGGG ACTTGGATCA ACCATCTCGT TACCTGGGGA TCTATTATAT TTTATTTTGT 1920  
 ATTTTCTTGG TTTTATGGAG GGATTCTCTG GCCATTTTGG GGCTCCAGAG ATATGTATTT 1980  
 TGTGTTTATT CAGCTCCTGT CAAGTGGTTC TGCTTGGTTT GCCATAATCC TCATGGTTGT 2040  
 TACATGTCTA TTTCTTGATA TCATAAGAA GGTCTTTGAC CGACACCTCC ACCCTACAAG 2100  
 85 TACTGAAAAG GTCACAGCTTA CTGAAAAGAA TGCAGGTATC AAGTGCTTGG ACTCCATGTG 2160  
 CTGTTTCCCG GAAGGAGAA GAGCGTGTGC ATCTGTTGGA AGAATGCTGG AACGAGTTAT 2220  
 AGGAAGATGT AGTCCAACCC ACATCAGCAG ATCATGGAGT GCATCGGATC CTTTCTATAC 2280  
 CAACGACAGG AGCATCTTGA CTCTCTCCAC AATGGACTCA TCTACTTGT TAAAGGGCAG 2340



5 TAGTACTTTG TGGGAGCCAG TTCACCTCCT TTCCTAAAAAT TCAGTGTGAT CACCCTGTTA 2400  
ATGGCCACAC TAGCTCTGAA ATTAATTTCC AAAATCTTTG TAGTAGTTCA TACCCACTCA 2460  
GAGTTATAAT GGCAAACAAA CAGAAAGCAT TAGTACAAGC CCCTCCCAAC ACCCTTAATT 2520  
TGAATCTGAA CATGTTAAAA TTTGAGAATA AAGAGACATT TTTTCATCTCT TTGTCTGGTT 2580  
TGTCCCTTGT GCTTATGGGA CTCCTAATGG CATTTCAGTC TGTGCTGAG GCCATTATAT 2640  
TTTAATATAA ATGTAGAAAA AAGAGAGAAA TCCTAGTAAA GAGTATTTTT TAGTATTAGC 2700  
TTGATTATTG ACTCTTCTAT TTAAATCTGC TTCTGTAAAT TATGCTGAAA GTTTGCCTTG 2760  
AGAAGCTCTAT TTTTTTATTA GAGTTATATT TAAAGCTTTT CATGGGAAAA GTTAATGTGA 2820  
ATACTGAGGA ATTTTGGTCC CTCAGTGACC TGTGTTGTTA ATTCATTAAT GCATTCTGAG 2880  
10 TTCACAGAGC AAATTAGGAG AATCATTTCC AACCATTATT TACTGCAGTA TGGGGAGTAA 2940  
ATTTATACCA ATTCCTCTAA CTGTACTGTA ACACAGCCTG TAAAGTTAGC CATATAAATG 3000  
CAAGGGTATA TCATATATAC AAATCAGGAA TCAGGTCCGT TCACCGAACT TCAAATTGAT 3060  
GTTTACTAAT ATTTTGTGTA CAGAGTATAA AGACCCTATA GTGGGTAAAT TAGATACTAT 3120  
TAGCATATTA TTAATTTAAT GTCTTTATCA TTGGATCTTT TGCATGCTTT AATCTGGTTA 3180  
15 ACATATTTAA ATTTGCTTTT TTTCTCTTTA CCTGAAGGCT CTGTGTATAG TATTTTCATGA 3240  
CATCGTTGTA CAGTTTAACT ATATCAATAA AAAGTTTGGA CAGTATTTAA ATATTGCAAA 3300  
TATGTTTAAAT TATACAAATC AGAATAGTAT GGGTAATTAA ATGAATACAA AAAGAAGAGC 3360  
CTCTTTCTGC AGCCGACTTA GACATGCTCT TCCCTTTCTA TAAGCTAGAT TTTAGAATAA 3420  
AGGGTTTCAG TTAATAATCT TATTTTCATC TATGTCATC TAACTTATAG CAAACTACCA 3480  
20 CAATACAGTG AGTTCGCCA GTGTCCAGT ACAAGGCATA TTTCAGGTGT GGCTGTGGAA 3540  
TGTAATAATG CTCAACTTGT ATCAGGTAAT GTTAGCAATA AATTAAATGC TAAGAATGAT 3600  
TAATCGGGTA CATGTTACTG TAATTAATCT ATTGCACTTC AAAACCTAAC TTCCATCCTG 3660  
AATTTATCAA GTAGTTCAGT ATTGTCAATT GTTTTTGTTT TATTGAAAAG TAATGTTGTC 3720  
TTAAGATTTA GAAGTGATTA TTAGCTTGAG AACTATTACC CAGCTCTAAG CAAATAATGA 3780  
25 TTGTATACAT ATTAAGATAA TGGTTAAATG CGGTTTACC AAGTTTTCCC TTGAAAATGT 3840  
AATTCCTTTA TGGAGATTTA TTGTGCAGCC CTTAGCTTCC TTTCCATTTT ATGAATATAA 3900  
GGCTTCTAGA ATTGACTGG CAGGGGAAAG AATGGTAGAG ACAGAAATTA AGACTTTATC 3960  
CTTGTCTGCT TGTAAGTAT TATTTTCTTG CTAATGTAAC ATTTGTCTGT TCCAGTGATG 4020  
TAAGGATATT AAGTTATTAA GCTAAATATT AATTTTCAAA AATAGTCCTT CTTTAACTTA 4080  
30 GATATTTTCA AGCTGGATTT AGGAAGATCT GTTATTCTGG AAGTACTAAA AAGAATAATA 4140  
CAACGTACAA TGTCTGCATT CACTAATTCA TGTTCCAGAA GAGGAAATAA TGAAGATATA 4200  
CTCAGTAGAG TACTAGGTGG GAGGATATGG AAATTTGCTC ATAAAATCTC TTATAAAACG 4260  
TGCATATAAC AAAATGACAC CCAGTAGGCC TGCATTACAT TTACATGACC GTGTTTATTT 4320  
GCCATCAAAAT AAAGTGAAGTA CTGACACCAG ACAAGACTC CAAAGTCATA AAATAGCCTA 4380  
35 TGACCAACTG CAGCAAGACA GGAGGTGAGC TCGCCTATAA TGGTGCTTAA AGTGTGATTG 4440  
ATGTAATTTT CTGTACTCAC CATTGGAAGT TAGTTAAGGA GAACCTTATT TTTTAAAAA 4500  
AAGTAAATGG CAACCACTAG TGTGCTCATC CTGAAGTGT ACTCCAAATC CACTCCGTTT 4560  
TTAAAGCAAA ATTATCTTGT GATTTTAAGA AAAGAGTTTT CTATTTATTT AAGAAAGTAA 4620  
CAATGCAGTC TGCAAGCTTT CAGTAGTTTT CTAGTGCTAT ATTCATCCTG TAAAACCTT 4680  
40 ACTACGTAAC CAGTAATCAC AAGGAAAGTG TCCCCTTTGC ATATTTCTTT AAAATTCCTT 4740  
CTTTGGAAAG TATGATGTTG ATAATTAAT TACCCTTATC TGCCAAAACC AGAGCAAAAT 4800  
GCTAAATACG TTATTGCTAA TCAGTGGTCT CAAATCGATT TGCTCCCTT TGCCTCGTCT 4860  
GAGGGCTGTA AGCCTGAAGA TAGTGGCAAG CACCAAGTCA GTTTCCAAA TTGCCCCTCA 4920  
45 GCTGCTTTAA GTGACTCAGC ACCCTGCCTC AGCTTCAGCA GGCGTAGGCT CACCCTGGGC 4980  
GGAGCAAAGT ATGGGCCAGG GAGAACTACA GCTACGAAGA CCTGCTGTCG AGTTGAGAAA 5040  
AGGGGAGAAT TTATGGTCTG AATTTTCTAA CTGTCTCTT TCTTGGGTCT AAAGCTCATA 5100  
ATACACAAAG GCTTCCAGAC CTGAGCCACA CCCAGGCCCT ATCCTGAACA GGAGACTAAA 5160  
CAGAGGCAAA TCAACCCTAG GAAATACTTG CATTCTGCCC TACGGTTAGT ACCAGGACTG 5220  
AGGTCAATTC TACTGGAAAA GATTGTGAGA TTGAACCTAT CTGATCGCTT GAGACTCCTA 5280  
50 ATAGGCAGGA GTCAAGGCCA CTAGAAATTA GACAGTTAAG AGCCAAAAGT TTTTAAAAA 5340  
TGCTACTCTG AAAAATCTCG TGAAGGCTGT AGGAAAAGGG AGAATCTTCC ATGTTGGTGT 5400  
TTTTCTCTGA AAGATCAGTT TGGGGTATGA TATAAGCAGG TATTAATAAA AATAACACAC 5460  
CAAAGAGTTA CGTAAACAT GTTTTATTAA TTTTGGTCCC CACGTACAGA CATTTTATTT 5520  
55 CTATTTTGAA ATGAGTTATC TATTTTCATA AAAGTAAAAC ACTATTAAAG TGCTGTTTTA 5580  
TGTGAATAAA CTTGAATGTT GTTCTTATAA AAAATAGATC ATAACCTCAT ATATGTTTGT 5640  
AATCATGGTA ATTTAGATTT TTATGAGGAA TGAGTATCTG GAAATATTGT AGCAATACTT 5700  
GGTTTAAAT TTTGGACCTG AGACACTGTG GCTGCTAAT GTAATCCTTT AAAAATTCTC 5760  
TGCATTGTCA GTAAATGTAG TATATTATTG TACAGCTACT CATAATTTTT TAAAGTTTAT 5820  
60 GAAGTTATAT TTATCAATAA AAAACTTTCC TATAT

Seq ID NO: 315 Protein sequence:  
Protein Accession #: XP\_087254

65 1 11 21 31 41 51  
| | | | | |  
MQFRECSING MKYQEINGRL VPEGTPDSS EGNLSYLSSL SHLNNLSHLT TSSSFRTSPE 60  
NETELIKEHD LFFKAVSLCH TVQISNVQTD CTGDGPWQSN LAPSQLEYA SSPDEKALVE 120  
AAARIGIVFI GNSEETMEVK TLGKLEERYKL LHILEFDSR RMSVIVQAP SGEKLLFAKG 180  
70 AEESILPKCI GGEIEKTRIH VDEFALKGLR TLCTIAYRKFT SKEYEEDKR IFEARTALQQ 240  
REEKLAADFQ FIEKDLILLG ATAVEDRLQD KVRTIEALR MAGIKVWVLT GDKHETAVSV 300  
SLSCGFHRT MNILELINQK SDSECAEQLR QLARRITEDH VIQGLVVDG TSLSLALREH 360  
EKLMEVCRN CSAVLCCRMA PLQKAKVIRL IKISPEKPIT LAVGDGANDV SMIQEAHVGI 420  
75 GIMGKEGRQA ARNSDYAIAR FKFLSKLLFV HGHFYIIRIA TLVQYFFYKN VCFITPQFLY 480  
QFYCLFSQQT LYDSVYLTLY NICFTSLPKL IYSLLEQHVD PHVLQNKPTL YRDISKNRLL 540  
SIKTLFYWTI LGFSHAFIF FGSYLLIGID TSLGNGQMF GNWTFGTLVF TVMVITVTVK 600  
MALETHFWTW INHLVTWGS IIFYVFSFLY GGILWPFLGS QNMYFVFIQL LSSGSAWFAI 660  
ILMVVTCFLF DIKKVFDRH LHPTSTKAQ LTETNAGIKC LDSMCCFPEG EAACASVGRM 720  
80 LERVIGRCSP THISRSWSAS DPFYTNDRSI LTLSTMDSS C

Seq ID NO: 316 DNA sequence  
Nucleic Acid Accession #: NM\_004473  
Coding sequence: 661..1791

85 1 11 21 31 41 51  
| | | | | |  
CTCGCCAGCG GTCCGCGGGG CTGGAGACCC ACGCCGTGGA GAGGACCAGC CTCAGGTCGC 60



5 CCGGCCTGGG CCGCGCCCC GACCTCGCTG CCCCCGCCTC GCCTCTCTGC CCGTGGCGCT 120  
 TACCGCCACC TTGGCCTCGG GGGCAGGGCA TGGGCGGGCC CCGCCAGATC GCCCAGCGCC 180  
 AGTACTAACT GCCCTCGCTC TGGCCTTCGA GCCCGAAGCC TCTTCTGCGC GCACAACCTA 240  
 GGCAGTAATC CTAAACTAGC GGGCACCACA GACCAGCTGC AGCCACCCCA ACCCAGGGAT 300  
 CACTTCCGGA CCCCTCGACC GCGCGGCACC AGCGCGCAAG GGACCCCTCA GCCGGAGACC 360  
 AGAGTCCAGT CCGGTCGCG AGGCCACCGC CGCTGCCCGC CTCGAGAAGC ACAACGCGGG 420  
 CTGAGCCGTC GGCTAGCGGG TCACTCCCAG GCCTCTGTCT GCACCGCGCC AGCCCCAGAC 480  
 CACGGACGCT GAGCCTCCAG CGCGCGCCAG CCTGGGCGCC TGGGCTCTCC GGGCCAGCCC 540  
 10 GCGACGATCC CCGAGCTCT CCGCAGAAGG GCCGAGCGTC CGTTCCGGGG ACGCCAGGCC 600  
 CGCCCCCGCC CCGGACAGC CGCGGGGATC CAGAGCCCGG GGGTGCGGGA CGCCCGCGCC 660  
 ATGACTGCGG AGAGCGGGCC GCGCGCGCCG CAGCCGAGG TGCTGGCTAC CGTGAAGGAA 720  
 GAGCGCGGCG AGACGGCAGC AGGGGCGGG GTCCCAGGGG AGGCCACGGG CCGCGGGGCG 780  
 GCGGGGCGGC GCCGAAGCG CCCCCTGCAG CGCGGGAAGC CGCCCTACAG CTACATCGCG 840  
 CTATCGCCA TGGCCATCGC GCACGCGCCG GAGCGCCGCC TCACGCTGGG CGGCATCTAC 900  
 15 AAGTTTATCA CCGAGCGCTT CCCCTTCTAC CGCGACAACC CCAAAAAGTG GCAGAACAGC 960  
 ATCCGCCACA ACCTCACACT CAACGACTGC TTCCTCAAGA TCCCGCGCGA GGCCGGCCGC 1020  
 CCGGGTAAGG GCAACTACTG GCGGCTCGAC CCAACGCGG AGGACATGTT CGAGAGCGGC 1080  
 AGCTTCTGTC GCCGCGCAA GCGCTTCAAG CGCTCGGACC TCTCCACCTA CCCGGCTTAC 1140  
 ATGCACGACG CGGCGGCTGC CGCAGCCGCC GCTGCCGCG CCGCCGCGGC CGCCGCGGCC 1200  
 20 GCGCCATCT TCCAGGCGC GGTGCGCGCC GCGCGCCCC CCTACCCGGG CGCCGTCTAT 1260  
 GCAGGCTACG CGCGCGGTC GCTGCGCGC CGGCTCCAG TCTACTACCC CGCGGCGTCG 1320  
 CCGCGCCCTT GCCGCGTCTT CCGCCTGGTT CCTGAGCGGC CGCTCAGCCC AGAGCTGGGG 1380  
 CCGCACCGT CGGGGCCCCG CGGCTCTTGC GCCTTTGCC TCCCGGCGC CCGCGCTACC 1440  
 25 ACCACCGGT ACCAGCCCGC AGGCTGCACC GGGGCCCCGG CGGCCAACCC CTCTGCCTAT 1500  
 GCGGCTGCCT ACGCCGGCCC CGACGGCGCG TACCCGCGAG CGCGCGGCG TGCGATCTTT 1560  
 GCGGCTGCTG GCCGCTGGC GGGACCCGCT TCGCCCCAG CGGGCGGCG CAGTGGCGGC 1620  
 GTGGAGACCA CGTGGAATCT CTACGGGCGC AGTGCGCCG GCCAGTTCGG AGCGCTGGGA 1680  
 GCCTGCTACA ACCCTGGCGG GCAGCTCGGA CCGGCGAGTG GCGCGCCTA CCATGCTCGC 1740  
 30 CATGCTGCCG CTTATCCCGG TGGGATAGAT CGGTTCGTGT CCGCCATGTG AGCCAGCGTA 1800  
 GGGACGAAAA CTCATAGACA CATCGGCTGT TCACACGTTT CCGCAACCT GAGAACGAAC 1860  
 AGGAATGGAG AGAGGACTCA ACTGGGACCC ACGTGGAAAA GACCGAGCAG GCCACAGAGG 1920  
 CTCGGTCTCC CCGCGCACAG CGTAGGCACC CTGTGTACTC TGTAACCGG AGGAGGTGGG 1980  
 GCGAGGACG CAGAGCCCTT GGAAGGACG AGGGACCTC GATGGAGCGA AGCCCTCAAA 2040  
 35 CGGGTCTGCT TCTGGCATTC TATCGGAGG GGTCTTGGC GGTAACGAG GGGCAGCGTA 2100  
 GTGTCAACAC CAGAGACCAG GATCCAAATT GTGGGGAATC AGTTTCAGCC TTCCATGTGC 2160  
 TGCCGGAAT CGGGCCTTTT TACGCGGTTT GTCTCTAGT GCCTTTAACT GCGTTACTAC 2220  
 AATAAAAGGC TGCGGACGCG CCTTTCTTCT TAAAGTGAGG AGGACAAATT TGCAAAAGAA 2280  
 ATAGGCTTTT CTTCTTTTCT AAATTGGAGA AATCTCTGCT CTGGTTGACC TGGGCTGGTT 2340  
 40 TTCCCTGTCT CTGAGAACTT GAGACCTAG TCCGAGTTGA ACTGTGCGTC AGCACTCCAG 2400  
 TCCCATCACC TGAACCTTCA GTCTCCCCGA TCTGTTACAC TAGAGGGCTG CAGGACTCTA 2460  
 TCCACCGCCC CCGGTTATC ATTACGGGCC CCATCATCTT GGATGCTGCC CTGCGTATTT 2520  
 GGCAGCAATG GTGGGCCACC CAGGGCCTCT GAGTAGCCAC CCAAAGCCTA GCCGCTGTTT 2580  
 TAGGGAACGG AAAAGAGTTC ATGGCCAAGC GTCTAACCTA AAGTCCCAGG ATTGGCTCCA 2640  
 45 GGCAGCAATT ATATCATAAC TTATTGAACT TTTGAGCAGG ACGTCTGGT AATTTTCATG 2700  
 CTGTTACTGC CCAGTCATAA ATCTGCTTTT CCATTATAAG GCAGAGAGAA GTACATTCTG 2760  
 TCATTTGTCC ACTGTTTCTT GTCATCACGC AGCCCTGGAC CCAAAGGGTG AACTAAAGTT 2820  
 TAAGGAGATG AGAGGATTCA AGGAGCCCGT TGGTGACGCC TTTCAGTAGC TGGGGAGGGC 2880  
 TCTTCCATCC CCAGCACCCC CTGCTACACC TCAGCAGCCT CCCCCATGCA AAAAGGAAAG 2940  
 50 AGAAAAATTA AGTTAGGGCA GTCAGTAAAG TGAGCTTTAG AAAGAACTG GAATTTTAAAC 3000  
 TTCATTTTGT ATCTTGCTTA AGTAGCAGGC TCACTAAAT TAGAGAAAGT CCAATAACTC 3060  
 TCCCCCTTTC CCTTGAGAAA TCTTTAAGTT TCGATTCTGG AGCAAAAAC TTCAGCATT 3120  
 AATATTTTCA AGGCTCCATT CACAGCTTTC AGATAAATG GAGTGTTCAG ATGGACTGTT 3180  
 TTAATAAAAA TCTTTGAGCA AGTGAGTTAT GGCAAGAGAA ACTCAGCCTC TTTCTGTATA 3240  
 55 AACTTAACAG GGAAGGGCTG GGGTGTGAAA AAGAAGATTG TATGAAAACC ATTGGTAATT 3300  
 TTTATTTTTT ATTTTGGGA CTGCACTATC CTGTTACAGA AGACATGTGA ACTTGGTTCA 3360  
 GTCCAAATGG GGATTTGTAT AAACCAGTGC TCTCCATTAG AAATATGGTG CAAGCCACAT 3420  
 ATGTAATTTT AAATATTCTA GTAGCCACAT TAATAAGTN AAAAGAAACA AAAAAAAAAA 3480  
 AA

Seq ID NO: 317 Protein sequence:  
 Protein Accession #: NP\_004464

65 1 11 21 31 41 51  
 | | | | |  
 FKHLTHYROI DTRANSCRIP TIONFACTOR TTFMTAESGP PPPQPEVLAT VKBERGETAA 60  
 GAGVPGEATG RGAGGRRRKR PLQRGKPPYS YIALIAMAIA HAPERRLTG GIYKFITERF 120  
 PFYRDNPCKW QNSIRHNLTL NDCFLKIPRE AGRPGKGNW ALDPNAEDMF ESGSFLRRRK 180  
 RFKRSDLSTY PAYMHDAAA AAAAAAAAAA AAAAAIFPGA VPAARPPYPG AVYAGYAPPS 240  
 70 LAAPPPVYYP AASPGPCRVF GLVPERPLSP ELGPAPSGPG GSCAFASAGA PATTTGYQPA 300  
 GCTGARPNP SAYAAAYAGP DGAYPQAGS AIFAAAGRLA GPASPPAGGS SGGVETTVDF 360  
 YGRTSPGQFG ALGACYNPGG QLGGASAGAY HARHAAAYPG GIDRFVSAM

Seq ID NO: 318 DNA sequence  
 Nucleic Acid Accession #: NM\_005688  
 Coding sequence: 126..4439

80 1 11 21 31 41 51  
 | | | | |  
 CCGGGCAGGT GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCTTGAGC 60  
 AGGGGCGCAG GAATTCTGAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCCGCTCAG 120  
 AGAAGATGAA GGATATCGAC ATAGGAAAAG AGTATATCAT CCCCAGTCCCT GGGTATAGAA 180  
 GTGTGAGGGA GAGAACCAGC ACTTCTGGGA CGCACAGAGA CCGTGAAGAT TCCAAGTTCA 240  
 85 GGAGAACTCG ACCGTTGGAA TGCCAGATG CTTTGGAAAC AGCAGCCCGA GCCGAGGGCC 300  
 TCTCTTTGA TGCCCTCCATG CATTCTCAGC TCAGAATCCT GGATGAGGAG CATCCCAAGG 360  
 GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACACCAGC 420  
 ACCCAGTGGA CAATGCTGGG CTTTTTTCCT GTATGACTTT TTCGTGGCTT TCTTCTCTGG 480



	CCCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAGAGC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
	AAGTTGGGCC	AGACGCTGCT	TCCCTGCGAA	GGGTGTGTGT	GATCTTCTGC	CGCACCAGGC	660
5	TCATCCTGTC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAGCCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTTAGT	GCTGGGCCTC	CTCCTGACGG	AAATCGTGGC	GTCTTGGTCC	CTTGCACTGA	840
	CTTGGGCATT	GAATTACCGA	ACCGGTGTCC	GCTTGCGGGG	GGCCATCCTA	ACCATGGCAT	900
	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
10	TTTGCTCCAA	CGATGGGCAG	AGAATGTTTG	AGGCAGCAGC	CGTTGGCAGC	CTGCTGGCTG	1020
	GAGGACCCGT	TGTTGCCATC	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCCTGGG	ATCAGCTGTT	TTTATCCTCT	TTTACCCAGC	AATGATGTTT	GCATCACGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CCGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTTCTC	1260
15	AGAGTGTTCA	AAAAATCCGC	GAGGAGGAGC	GTCCGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGTATCAC	TGTGGGTGTG	GCTCCCATTG	TGGTGGTGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTCATAT	GACCCCTGGG	TTGATCTGTA	CAGCAGCACA	GGCTTTCACA	GTGGTGACAG	1440
	TCTTCAATT	CATGACTTTT	GCTTTGAAAG	TAACACCGTT	TTTCAATAAG	TCCCTCTCAG	1500
	AAGCCTCAGT	GGCTGTTGAC	AGATTTAAGA	GTTTGTCTCT	AATGGAAGAG	GTTTACATGA	1560
20	TAAAGAACAA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAAAAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCACTCCAGT	ATCCAGAACT	CGCCCAAGCT	GACCCCAAAA	ATGAAAAAAG	1680
	ACAAGAGGGC	TTCCAGGGGC	AAGAAAGAGA	AGGTGAGGCA	GCTGCAGCGC	ACTGAGCATC	1740
	AGGCGGTGCT	GGCAGAGCAG	AAAGGCCACC	TCCTCCTGGA	CAGTGACGAG	CGGCCAGTGC	1800
	CCGAAGAGGA	AGAAGGCAAG	CACATCCACC	TGGGCCACCT	GCGCTTACAG	AGGACACTGC	1860
25	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AACTGGTTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCAT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACCTTC	GCTTATGTGG	CCCAGCAGGC	CTGGATCCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTG	CCTGAGGCCT	GACCTGGCCA	TTCTTCCAG	CAGCGACCTG	ACGGAGATTG	2160
30	GAGAGCGAGG	AGCCAACCTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCGGGCCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTT	CAGTGCCTTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGATCTTC	ATGAAAGAGG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCATGAGAG	AACTGATGAA	TTTAAATGGT	GACTATGCTA	2460
35	CCATTTTAA	TAACCTGTTG	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAAAG	2520
	AAACCAGTGG	TTCACAGAAG	AAGTCACAAG	ACAAGGGTCC	TAAAACAGGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTTCAGTGC	CTGGTCAGTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
	TCCTGGTTAT	TATGGCCCTT	TTTATGCTGA	ATGTAGGCAG	CACCGCCTTC	AGCACCTGGT	2760
40	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAAACACC	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCCTGA	AAGCCATTCT	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGCG	AGCTTCTCTC	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTCAAGGCC	3000
	CTATGAAGTT	TTTTGACACG	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
45	TGGATGAAGT	TGACGTGCGG	CTGCCGTTCC	AGGCCGAGAT	GTTTATCCAG	AACGTTATCC	3120
	TGGTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCTT	GTGGCAGTGG	3180
	GGCCCCCTGT	CATCCTCTTT	TCAGTCCTGC	ACATTTCTCT	CAGGGTCCTG	ATTCGGGAGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCAGTCAC	CTTCTCTCTC	CCACATCAGC	TCCAGCATA	3300
	AGGGCCTTGC	CACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTTCTGCAC	AGATACCAGG	3360
50	AGCTGCTGGA	TGACAACCAA	GCTCCTTTTT	TTTTGTCTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACCAC	CACGGGGCTG	ATGATCGTTC	3480
	TTATGCACGG	GCAGATTCCC	CCAGCCTATG	CGGGTCTCGC	CATCTCTTAT	GCTGTCCAGT	3540
	TAACGGGGCT	GTTCCAGTTT	ACGGTCAGAG	TGCTCATCTG	GACAGAAGCT	CGATTACCTT	3600
	CGGTGGGAGG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAAGCACC	GCCAGAATTA	3660
55	AGAACAAGGC	TCCCTCCCCC	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGCAG	3720
	AGATGAGGTA	CCGAGAAAAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAAC	3780
	CTAAAGAGAA	GATTGGCATT	GTGGGGCGGA	CAGGATCAGG	GAAATCCTCG	CTGGGGATGG	3840
	CCCTCTTCCG	TCTGGTGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAAATCA	3900
	GTGATATTGG	CCTTGCCGAC	CTCCGAAGCA	AACTCTCTAT	CATTCTCTCA	GAGCCGGTGC	3960
60	TGTTTCAGTG	TCAATTTTGG	TCAAAATTTG	ACCCCTTCAA	CCAGTACACT	GAAGACCAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TTGAATCTGA	AGTGATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
	CCATGGACAC	AGAGACAGAC	TTATTGATTG	AAGAGACCAT	CCGAGAAGCA	TTTGCAGACT	4260
65	GTACCATGCT	GACCATTGCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTTCTG	TCCAACGACA	4380
	GTTCCCGATT	CTATGCCATG	TTTGCTGCTG	CAGAGAACAA	GGTCGCTGTC	AAGGGCTGAC	4440
	TCCTCCCTGT	TGACGAAGTC	TCTTTTCTTT	AGAGCATTGC	CATTCCCTGC	CTGGGGCGGG	4500
	CCCCTCATCG	CGTCTCTCTA	CCGAAACCTT	GCCTTTCTCG	ATTTTATCTT	TCCACAGCA	4560
70	GTTCCGGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTGTATTT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTTTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAAGGTC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
	TATTAAAATA	AGCACTGTGC	TAATAACAGT	GCATATTCCCT	TTCTATCATT	TTTGTACAGT	4860
75	TTGCTGTACT	AGAGATCTGG	TTTTGTCTAT	AGACTGTAGG	AAGAGTAGCA	TTTCATTCTT	4920
	CTCTAGCTGG	TGGTTTCACG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCCTCTGCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
	GAGACGGGTG	GGCGGCTGGA	GACCATGCAG	AGCGCCGTGA	GTTCTCAGGG	CTCCTGCCTT	5100
	CTGTCTCTGT	GTCACCTACT	GTTTCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCTT	5160
80	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTCTGCT	TCTTCTTTTT	GCTGTTGTTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACCTGC	TCAGGTTTCT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340
	GTTGGTTCCA	AGCCCTGGAG	CCAACCTGCT	CTTTTTGAGG	TGGCACTTTT	TCATTTGCCT	5400
	ATTCACACAC	CTCCACAGTT	CAGTGGCAGG	GCTCAGGATT	TCGTGGGTCT	GTTTTCTTTT	5460
85	CTCACCGCAG	TGCTCGCACA	GTCTCTCTCT	CTCTCTCCCC	TCAAAGTCTG	CAACTTTAAG	5520
	CAGCTCTTGC	TAATCAGTGT	CTCACACTGG	CGTAGAAGTT	TTTGTACTGT	AAAGAGACCT	5580
	ACCTCAGGTT	GCTGGTTGCT	GTGTGGTTTG	GTGTGTCTCC	GCAAAACCCC	TTTGTGCTGT	5640
	GGGGCTGGTA	GCTCAGGTGG	GCGTGGTCAC	TGCTGTCTAT	AGTTGAATGG	TCAGCGTTGC	5700



ATGTCGTGAC CAACTAGACA TTCTGTGCGC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760  
CAAAAATCTG AAAATGTGAA TAAATATTATT TTGGATTTTG TAAAAAATAA AAAAAAATAA 5820  
AAAAAATAA AAAAAATAA

5

Seq ID NO: 319 Protein sequence:  
Protein Accession #: NP\_005679

10 1 11 21 31 41 51  
| | | | | |  
MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRF TRPLECQDAL ETAARAEGLS 60  
LDASMHSQLR ILDEEHPKGK YHHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSSLAR 120  
VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQEELNEV GPDAASLRRV VWIFCRTRLI 180  
LSIVCLMITQ LAGFSGPAFM VKHLLLEYTQA TESNLQYSL LVLGLLLTEI VRSWSLALTW 240  
15 ALNYRTGVRL RGAILTMAFK KILKLKNIKE KSLGELINIC SNDGQRMFEA AAVGSLLAGG 300  
PVVAILGMIY NVIILGPTGF LGSVAVFILFY PAMMFASRLT AYFRRKCVAA TDERVQKMNE 360  
VLTYIKFIKM YAWVKAFSQS VQKIREEEERR ILEKAGYFQG ITVGVAPIVV VIASVVTFSV 420  
HMTLGFDLTA AQAFVTVVTF NSMTFALKVT PFSVKSLSEA SVAVDRFKSL FLMEEVHMIK 480  
NKPASPHIKI EMKNATLAWD SSHSSIQNSP KLTPKMKKDK RASRGKKEKV RQLQRTHEQA 540  
20 VLAEQKGHL LLSDERPSPE EEEKGHIHLG HLRLQRTLHS IDLEIQEGKL VGICGSVSGS 600  
KTSLSAILG QMTLLEGSIA ISGTFAYVAQ QAWILNATLR DNILFGKEYD EERYNSVLNS 660  
CCLRPDLAIL PSSDLTEIGE RGNLSSGGQR QRSILARALY SDRSIYILDD PLSALDAHVG 720  
NHIFNSAIRK HLKSKTVLFV THQLQYLVDC DEVIFMKEGC ITERGTHEEL MNLNGDYATI 780  
FNNLLLGETP PVEINSKKET SGSQKKSQDK GPKTGSVKKE KAVKPEEGQL VQLEEKQGS 840  
25 VPWSVYGVYI QAAGGPLAFL VIMALFMLNV GSTAFSTWWL SYWIKQSGSN TTVTRGNETS 900  
VSDSMKDNPH MQYASIYAL SMAVMLILKA IRGVVFKVGT LRASSRLHDE LFRRILRSPM 960  
KFFDTTPTGR ILNRFSKDM D EVDVRLPFQA EMFIQNVILV FFCVGMIAV FPWFLVAVGP 1020  
LVILFSLVHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QEFLHRYQEL 1080  
LDDNQAPFL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HGQIPPAYAG LAISYAVQLT 1140  
30 GLFQFTVRLA SETEARFTSV ERINHYIKTL SLEAPARIKN KAPSPDWPQE GEVTFENAEM 1200  
RYRENLPVLV KKVSTFIKPK EKIGIVGRIG SGKSSLGML FRLVELSGGC IKIDGVRISD 1260  
IGLADLRSLK SIIPQEPVLF SGTVRSNLDN FNQYTEDQIW DALERTHMKE CIAQLPLKLE 1320  
SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIIQE TIREAFADCT 1380  
35 MLTIAHRLHT VLGS DRIMVL AQGQVVEFDT PSVLLSNDSS RPYAMFAAAE NKVAVKKG

Seq ID NO: 320 DNA sequence  
Nucleic Acid Accession #: AK022089.1  
Coding sequence: 181-1488

40 1 11 21 31 41 51  
| | | | | |  
AGCAGTTGCA CAACTTCCAG CAACTTTCTC AGCCGGCTAC TAATGAGCTG AAAGCCAGGA 60  
ACATCCGAGG AGAAGAGAAA GCTTCCAGCC CTCCTCCCTT CACCCTGGAA ATCCAGACAC 120  
45 CCCACCCCC ACCCTCAGAT CACTTTAAGA TAATTTCTTT ATTCGTTTGC CCGACAGACC 180  
ATGGCTCCCT TTGGAAGAAA CTTGCTAAAG ACTCGGCATA AAAACAGATC TCCAATAAAA 240  
GACATGGATT CAGAAGAGAA GGAAATTGTG GTTTGGGTTT GCCAAGAAGA GAAGCTTGTC 300  
TGTGGGCTGA CTAAACGCAC CACCTCTGCT GATGTCATCC AGGCTTTGCT TGAGGAACAT 360  
GAGGCTACGT TTGGAGAGAA ACGATTTCTT CTGGGGAAGC CCAGTGATTA CTGCATCATA 420  
50 GAGAAGTGGG GAGGCTCCGA AAGGGTTCTT CCTCCACTAA CTAGAATCCT GAAGCTTTGG 480  
AAAGCGTGGG GAGATGAGCA GCCCAATATG CAATTTGTTT TGGTTAAAGC AGATGCTTTT 540  
CTTCCAGTTC CTTTGTGGCG GACAGCTGAA GCCAAATTAG TGCAAAACAC AGAAAAATTG 600  
TGGGAGCTCA GCCCAGCAA ACTTTACCAC CAGATAAACA AAAAAGAATA 660  
GTCAGGAAAA CTTTCCGGA ACTGGCTAAA ATTAAGCAGG ACACAGTTTC TCATGATCGA 720  
55 GATAATATGG AGACATTAGT TCATCTGATC ATTTCCCAGG ACCATACTAT TCATCAGCAA 780  
GTCAAGAGAA TGAAAGAGCT GGATCTGGAA ATTGAAAAGT GTGAAGCTAA GTTCCATCTT 840  
GATCGAGTAG AAAATGATGG AGAAAACTAT GTTCAGGATG CATATTTAAT GCCCAGTTTC 900  
AGTGAAGTTG AGCAAAATCT AGACTTGCAG TATGAGGAAA ACCAGACTCT GGAGGACCTG 960  
AGCGAAAGTG ATGGAATTGA ACAGCTGGA GAACGACTGA AATATTACCG AATACTCATT 1020  
60 GATAAGCTCT CTGCTGAAAT AGAAAAAGAG GTAAAAAGTG TTTGCATTGA TATAAATGAA 1080  
GATGCGGAAG GGGAGCTGCA AAGTGAAGTG GAAAGCTCTA ATTTAGAGAG TGTAAAGTGT 1140  
GATTTGGAGA AAAGCATGAA AGCTGGTTTG AAAATTCAC CTCAATTGAG TGGCATCCAG 1200  
AAAGAGATTA AATACAGTGA CTCATTGCTT CAGATGAAAG CAAAAGAATA TGAACCTCTG 1260  
GCCAAGGAAT TCAATTCAC TCACATTAGC AACAAAGATG GGTGCCAGTT AAAGGAAAAC 1320  
65 AGAGCGAAGG AATCTGAGGT TCCAGTAGC AATGGGGAGA TTCCTCCCTT TACTCAAAGA 1380  
GTATTTAGCA ATTACACAAA TGACACAGAG TCGGACACTG GTATCAGTTC TAACCACAGT 1440  
CAGGACTCCG AAACAACAGT AGGAGATGTG GTGCTGTTGT CAACATAGTT CCAATGGCTC 1500  
CTTTCTGACC TGCTTTCATG TTTTAATGTT TGTTTAATTT AATAGGAAAC CTCATTTTAA 1560  
70 ATATAACACT CAAAAAATG TAAATCATAT TGTAGTATTC AATAGTTAAT AAAAATCTGA 1620  
GAAATGTGTT GTTTCTG

Seq ID NO: 321 Protein sequence:  
Protein Accession #: NP\_005438.1

75 1 11 21 31 41 51  
| | | | | |  
MAPFGRNLLK TRHKNRSPK DMDSEEKEIV VWVCQEEKLV CGLTKRTTSA DVIQALLLEH 60  
EATFGEKRFL LGKPSDYCII EKWRGSESVL PPLTRILKLW KAWGDEQPNM QFVLVKADAF 120  
80 LPVPLWRTAE AKLVQNTKEL WELSPANMYK TLPPDKQKRI VRKTFRKLAK IKQDTVSHDR 180  
DNMETLVHLI ISQDHTIHQQ VKRMKELDL IEKCEAKFHL DRVENDGENY VQDAYLMPSF 240  
SEVEQNLDLQ YEENQTLDEL SESDGIEQLE ERLKYRILI DKLSAEIEKE VKSVCIDINE 300  
DAEGEAASEL ESSNLESVKC DLEKSMKAGL KIHSHLSGIQ KEIKYSDSL QMKAKEYELL 360  
AKEFNSLHIS NKDGCQLKEN RAKESEVPSS NGEIPPFTQR VFSNYTNDTD SDTGISSNHS 420  
85 QDSETTVGDV VLLST

Seq ID NO: 322 DNA sequence  
Nucleic Acid Accession #: NM\_030920.1



Coding sequence: 317-1123

	1	11	21	31	41	51	
5							
	AGCATTGAAG	GGGAAGGAAC	TGCGGGTGTG	GTGTGTGTAT	GTGTGTGTGT	ATGTGTGTGC	60
	GGCGCGTGCG	TGCGTGTGTG	TGCGCGCGCT	AGTGTGTGGA	CAAGGAGGTG	GGGGCAGCTG	120
	AGTTAGAGTC	CCAACCTCTG	GACTCCATTT	GCTATTCTCT	TCTTTCTCCC	CCACACCTAT	180
	CTGGTGGTGG	TAGTGGGCGT	TTATATTTGC	GTTCCCTTTT	ATTCATTTCT	AAATCTCTTA	240
10	AAAATTTTGG	GTTGGGGGTA	TTGGGGAAGG	CAGGAAAGGG	AAAAGGAGAG	TAGTAGCTGA	300
	AGAGCAAGAG	GAGGACATGG	AGATGAAGAA	GAAGATTAAC	CTGGAGTTAA	GGAACAGATC	360
	CCCGGAGGAG	GTGACAGAGT	TAGTCCTTGA	TAATTGCCTG	TGTGTCAATG	GGGAAATTGA	420
	AGGCCTGAAT	GATACTTTCA	AAGAACTAGA	ATTTCTGAGT	ATGGCTAATG	TGGAACATAAG	480
	TTTCGCTGGCC	CGGCTTCCCA	GCTTAAATAA	ACTTCGAAAA	TTGGAGCTTA	GTGATAATAT	540
15	AATTTCTGGA	GGCTTGGAAG	TCCTGGCAGA	GAAATGTCCA	AATCTTACCT	ACCTCAATCT	600
	GAGTGGAAAC	AAAATAAAAG	ATCTCAGTAC	AGTAGAAGCT	CTGCAAAATC	TTAAAAATTT	660
	GAAAAGTCTT	GACCTGTTTA	ACTGTGAGAT	CACAAACCTG	GAAGATTATA	GAGAAAGTAT	720
	TTTTGAACTA	CTGCAGCAAA	TCACATACTT	AGATGGATTT	GATCAGGAGG	ATAATGAAGC	780
	GCCGGACTCT	GAAGAGGAGG	ATGATGAGGA	TGGAGATGAA	GATGATGAAG	AGGAAGAGGA	840
20	AAATGAAGCT	GGTCCACCGG	AAGGATATGA	GGAAGAGGAG	GAGGAAGAGG	AAGAGGAGGA	900
	TGAGGATGAG	GATGAAGATG	AAGATGAAGC	AGGTTGAGAG	TTGGGAGAGG	GAGAAGAGGA	960
	AGTGGGCTCT	TCATACTTAA	TGAAAGAAGA	AATTCAGGAT	GAAGAAGATG	ATGATGACTA	1020
	TGTTGAAGAA	GGGGAAGAAG	AGGAAGAAGA	GGAAGAAGGA	GGTCTTCGAG	GGGAGAAGAG	1080
	GAAACGAGAT	GCTGAAGACG	ATGGAGAGGA	AGAAGATGAC	TAGATCATTC	TAAGACCAGA	1140
25	TTCTCTAATG	TTTCTGGGTG	TGCAATAGAG	TGATCACATC	TTTGTTCCTT	CATGTACGAT	1200
	AGCTATCCCT	ACAGAAGATA	ATGTGTAAC	TTTTATAGGA	AAAGTGTGGT	TTTACTATTT	1260
	TTGCCTTATC	ATTCCAAATA	AGAAGTAGTC	TGTTAATGAT	CATATTGTAT	GTAGAGAAAA	1320
	ATTTTCATTG	ACTCCCATTG	TGGAATTCCC	TAGCAATTTA	TTTAGACTTA	ATTTTTTAAA	1380
	TTCAAGCTTA	CTGTATTAGT	CATTTTATAG	CCATAATTAA	AACATGATCA	CTTTTAAACA	1440
30	GGTGTAGTAT	GGTGCATTTC	ATTCCCTTAT	TATAGATTAA	CTGAAATTAC	AGTTTGCTAT	1500
	AATATAAAAT	GACAATAGTC	TCTTGAGTGG	TAAGTTGGTT	ATTTTTTTAG	AGGTGATCCA	1560
	GGAATCTTTA	GTTTGAAGGC	AGTTACCTTT	TTTTTTTTTT	TTTTTTTTTG	ACTAAGAGTG	1620
	TTTGGTTGCT	TTTTTGTGAC	AAGTAACTTG	GAAAATAGAA	GCAGAATAGT	AAAGGTTCTA	1680
	TTCAGCAACA	TAGTTCATGG	ATTTTGTGGA	GGTTCATTTC	AGTAATATGG	TTCATGGATT	1740
35	TAGTGGTGAC	TGATAAGATT	TTATTTTGA	AGGAAAAATT	GCTTATACTA	AGTCCAGAGA	1800
	CATGCAGGTG	AGCCCTTTTG	TCAGGCTGCA	AATCATGACA	TGCCGATGGT	TGTTTATTTT	1860
	GTTTTTAGGT	GTGCATTCTT	TTTCTTCTTA	GCAATTCCTT	TATGATCACC	TTCCCTTCTT	1920
	GTTTCACTCC	CTCCCGCTCT	CTCAAAAGGA	ACTTGGGAAA	CTTGTGAAAC	CCAGGAAAAC	1980
	CTTTAGTCTT	ATACCTCAAC	TACGTTTCAG	TCCTGTCTGG	GTTTTAAATA	AGTGAAGTAG	2040
40	AAGAAATTGA	GTATTTTCTG	ACATAAGAAT	ATATTATCAA	TACAGTTTFA	TGCAGTAAGC	2100
	TCTCCTTACC	ATAAATGTTT	CTTGTTTGAC	AACATCTAAG	ACAATATTAG	TGGGATGAAG	2160
	AAAGAAAAGC	AGGGGTGCTT	TTGGAAGCAG	TGTTAGTGTT	CCTCAAAAGT	CGGAACAATT	2220
	GCCTGTTGAT	ATATTAATAA	GACATTAAAG	TCAAATTTTA	ATGTTGGCCT	CTCAAATGAT	2280
	TTGGATACCA	CTCTGCAAAG	TATTTCTAAC	CTTTAATTCC	CAGTTTFAAA	ACAGATATAA	2340
45	TAATAGCATT	TAATTGGAAT	ATACTAGGCA	GCTGGAAAAG	TATTTGAAAC	TAAATTGACA	2400
	TTAAAATTAA	GATTTGTTTT	CAAGTGGATG	TCCATTAAAA	GTAGAAAAAT	ATTTGGGATA	2460
	AGTGAAGTGT	TGTTTCCCTA	CATGGCTACT	AAATAAAATA	TAATGAGTAT	ACAAGTATAT	2520
	CTCCTCTTTT	GCTATGGAGG	CTCCATGTTT	AAGGCAATGG	CTTTTAAAT	CTTGGCTATC	2580
	TAAAATTTTT	TCCCTTTGTT	TTGAATATTT	GTAAGTTTTT	AAGAAGTTAG	TGTCAGCAAA	2640
50	TTAATTGAAG	TTATGCTTCT	ATACTGGGAC	ATATTTAAAT	ACTGAGTATA	GTACTGCTGC	2700
	TACTGCTTCT	ACAATGTAAA	ATGTATGACT	TGGTGTTTTA	AAGTAAAAAT	TATGATGTTA	2760
	CTTGTGGAGA	AACATAAAAT	GTGTGACAAC	TGACCGAAAG	AAAACCCCTG	GGGATAAGTT	2820
	TAGTGAAGGG	ATTGGAATCC	CCAAAAAGAT	AACATTTTTC	TTCTGCTTTT	AAAAACTGAA	2880
	ATTCCCTGTT	CTAGTTCCTA	ACAATTTCTA	TTACATACTA	TGCCAGATTA	CAAAATACTT	2940
55	ATTTTAAAAA	TGAAATCTAT	ATATTGACTT	TCTTATCAAT	CATCTTACTG	TGCAATCAAA	3000
	ATTAGAGTAC	TTTGGTTTGA	AAACAACACT	TAGAGCCTCC	AGATAACTTT	TAAGACTTAT	3060
	TTAGCTTTGT	GGGTGGTATT	TTCATGCAAA	TAAGTAAGGG	TGGGTTTAT	ATTTTGTAGA	3120
	AGTTTTCGGT	CCTATTTTAA	TGCTCTTTGT	ATGGCAGTAT	GTATATATTG	TGTTAAGTTC	3180
60	CTCAAGAATC	TCCTTAAAAA	CTTTGAAGTT	AATACTTTTG	TGCAACTGTG	TTTTGAATAA	3240
	AGCCATGACA	GTGTTAAAAA	CAAAC				

Seq ID NO: 323 Protein sequence:  
Protein Accession #: NP\_112182.1

65	1	11	21	31	41	51	
	MEMKKKINLE	LRNRSPEEVT	ELVLDNCLCV	NGEIEGLNDT	FKELEFLSMA	NVELSSLARL	60
	PSLNKLRKLE	LSDNIISGGL	EVLAEKCPNL	TYLNLSGNKI	KDLSTVEALQ	NLKNLKSLLD	120
	FNCEITNLED	YRESIFELLQ	QITYLDGFDQ	EDNEAPDSEE	EDDEDGDEDD	EEEEENEAGP	180
70	PEGYEEEEEE	EEEEDEDEDE	DEDEAGSELG	EGEEVEGLSY	LMKEEIQDEE	DDDDYVEEGE	240
	EEEEEEEGGL	RGEKRKRDAE	DDGEEEDD				

Seq ID NO: 324 DNA sequence  
Nucleic Acid Accession #: NM\_003812  
Coding sequence: 224..2722

75	1	11	21	31	41	51	
	TCTCTGCGT	CCCGCCCCGG	GAGTGGCTGC	GAGGCTAGGC	GAGCCGGGAA	AGGGGGCGCC	60
	GCCCAGCCCC	GAGCCCCGCG	CCCCGTGCCC	CGAGCCCCGA	GCCCCCTGCC	CGCGGCGGCA	120
80	CCATGCGCGC	CGAGCCGGCG	TGACCGGCTC	CGCCCGCGGC	CGCCCCGCGC	CTAGCCCCGC	180
	GCTCTGCGCG	GCCACACGGA	GCGGCGCCCG	GGAGCTATGA	GCCATGAAGC	CGCCCCGCGC	240
	CAGCTGCGCG	CAGCCGCCCC	TGGCGGGCTG	CAGCCTTGCC	GGCGCTTCCT	GCGGCCCCCA	300
	ACGCGGCCCC	GCCGGCTCGG	TGCCTGCCAG	CGCCCCGGCC	CGCACGCCGC	CCTGCCGCCT	360
	GCTTCTCGTC	CTTCTCCTGC	TGCCTCCGCT	CGCCGCTCG	TCCCGGCCCC	GCGCCTGGGG	420
85	GGCTGCTGCG	CCCAGCGCTC	CGCATTGGA	TGAAACTGCA	GAAAAAATT	TGGGAGTCCT	480
	GGCAGATGAA	GACAATACAT	TGCAACAGAA	TAGCAGCAGT	AATATCAGTT	ACAGCAATGC	540
	AATGCAGAAA	GAAATCACAC	TGCCTTCAAG	ACTCATATAT	TACATCAACC	AAGACTCGGA	600



	AAGCCCTTAT	CACGTTCTTG	ACACAAAGGC	AAGACACCAG	CAAAAACATA	ATAAGGCTGT	660
	CCATCTGGCC	CAGGCAAGCT	TCCAGATTGA	AGCCTTCGGC	TCCAAATTCA	TTCTTGACCT	720
	CATACTGAAC	AATGGTTTGT	TGTCTTCTGA	TTATGTGGAG	ATTCACACG	AAAATGGGAA	780
5	ACCACAGTAC	TCTAAGGGTG	GAGAGCACTG	TTACTACCAT	GGAAGCATCA	GAGGCGTCAA	840
	AGACTCCAAG	GTGGCTCTGT	CAACCTGCAA	TGGACTTCAT	GGCATGTTTG	AAGATGATAC	900
	CTTCGTGTAT	ATGATAGAGC	CACTAGAGCT	GGTTTCATGAT	GAGAAAAGCA	CAGGTCGACC	960
	ACATATAATC	CAGAAAACCT	TGGCAGGACA	GTATTCTAAG	CAAATGAAGA	ATCTCACTAT	1020
	GGAAAGAGGT	GACCAGTGGC	CCTTTCTCTC	TGAATTACAG	TGGTTGAAAA	GAAGGAAGAG	1080
10	AGCAGTGAAT	CCATCACGTG	GTATATTTGA	AGAAATGAAA	TATTTGGAAC	TTATGATTGT	1140
	TAATGATCAC	AAAACGTATA	AGAAGCATCG	CTCTTCTCAT	GCACATACCA	ACAACTTTGC	1200
	AAAGTCCGTG	GTCAACCTTG	TGGATTCTAT	TTACAAGGAG	CAGCTCAACA	CCAGGGTTGT	1260
	CCTGGTGGCT	GTAGAGACCT	GGACTGAGAA	GGATCAGATT	GACATCACCA	CCAACCTGT	1320
	GCAGATGCTC	CATGAGTTCT	CAAAATACCG	GCAGCGCATT	AAGCAGCATG	CTGATGCTGT	1380
15	GCACCTCATC	TCGCGGGTGA	CATTTCACTA	TAAGAGAAGC	AGTCTGAGTT	ACTTTGGAGG	1440
	TGTCTGTTCT	CGCACAAAGAG	GAGTTGGTGT	GAATGAGTAT	GGTCTTCCAA	TGGCAGTGGC	1500
	ACAAGTATTA	TCGCAGAGCC	TGGCTCAAAA	CCTTGGGAATC	CAATGGGAAC	CTTCTAGCAG	1560
	AAAGCCAAAA	TGTGACTGCA	CAGAATCCTG	GGGTGGCTGC	ATCATGGAGG	AAACAGGGGT	1620
	GTCCCATTCT	CGAAAATTTT	CAAAGTGCAG	CATTTTGGAG	TATAGAGACT	TTTTACAGAG	1680
	AGGAGGTGGA	GCCTGCCTTT	TCAACAGGCT	AACAAAGCTA	TTTGAGCCCA	CGGAATGTGG	1740
20	AAATGGATAC	GTGGAAGCTG	GGGAGAGGTC	GATTGTGGT	TTTCATGTGG	AATGCTATGG	1800
	ATTATGCTGT	AAGAAATGTT	CCCTCTCCAA	CGGGGCTCAC	TGCAGCGACG	GGCCCTGCTG	1860
	TAACAATACC	TCATGTCTTT	TCAGCCACG	AGGGTATGAA	TGCCGGGATG	CTGTGAACGA	1920
	GTGTGATATT	ACTGAATATT	GTACTGGAGA	CTCTGGTCAG	TGCCACCAA	ATCTTCATAA	1980
25	GCAAGACGGA	TATGCATGCA	ATCAAAATCA	GGGCCGCTGC	TACAATGGCG	AGTGCAAGAC	2040
	CAGAGACAAC	CAGTGTCACT	ACATCTGGG	AACAAAGGCT	GCAGGGTCTG	ACAAGTTCTG	2100
	CTATGAAAAG	CTGAATACAG	AAGGCACTGA	GAAGGGAAAC	TGCGGGAAGG	ATGGAGACCG	2160
	GTGGATTGAG	TGCAGCAAAC	ATGATGTGTT	CTGTGGATTG	TTACTCTGTA	CCAATCTTAC	2220
	TCGAGCTCCA	CGTATTGGTC	AACTTCAGGG	TGAGATCATT	CCAATCTCCT	TCTACCATCA	2280
30	AGGCCGGGTG	ATTGACTGCA	GTGGTGCCCA	TGTAGTTTTA	GATGATGATA	CGGATGTGGG	2340
	CTATGTAGAA	GATGGAACGC	CATGTGGCCC	GTCTATGATG	TGTTTAGATC	GGAAGTGCCT	2400
	ACAAATTCAA	GCCCTAAATA	TGAGCAGCTG	TCCACTCGAT	TCCAAGGGTA	AAGTCTGTTC	2460
	GGGCCATGGG	GTGTGTAGTA	ATGAAGCCAG	ALSTCNLTGT	GATTTACCT	GGGCAGGGAC	2520
	AGATTGCAGT	ATCCGGGATC	CAGTTAGGAA	CCTTCACCCC	CCCAAGGATG	AAGGACCCAA	2580
35	GGGTCTAGT	GCCACCAATC	TCATAATAGG	CTCCATCGCT	GGTGCCATCC	TGGTAGCAGC	2640
	TATTGTCTTT	GGGGGCACAG	GCTGGGGATT	TAAAAATGTC	AAGAAGAGAA	GGTTCGATCC	2700
	TACTCAGCAA	GGCCCCATCT	GAATCAGCTG	CGCTGGATGG	ACACCGCCTT	GCACTGTTGG	2760
	ATTCTGGGTA	TGACATACTC	GCAGCAGTGT	TACTGGAACT	ATTAAGTTTG	TAAACAAAAC	2820
	CTTTGGGTGG	TAATGACTAC	GGAGCTAAAG	TTGGGGTGAC	AAGGATGGGG	TAAAAGAAAA	2880
40	CTGTCTCTTT	TGGAATAAAT	GTCAAAAGAA	ACCTTTCACC	ACCTGTCACT	AAACGGGGGA	2940
	GGGGGCAAAA	GACCATGCTA	TAAAAAGAAC	TGTTCCAGAA	TCTTTTTTTT	TCCCTAATGG	3000
	ACGAAGGAAC	AACACACACA	CAAAAATTAA	ATGCAATAAA	GGAATCATT	AAAA	

Seq ID NO: 325 Protein sequence:

Protein Accession #: NP\_003803

45	1	11	21	31	41	51	
	MKPPGSSSRQ	PPLAGCSLAG	ASCGPQRGPA	GSVPASAPAR	TPPCRLLLV	LLLPPLAASS	60
	RPRAWGAAAP	SAPHWNETAE	KNLGVLADE	NTLQONSSSN	ISYSNAMQKE	ITLPSRLIYY	120
50	INQDSESPYH	VLDTKARHQQ	KHNKAVHLAQ	ASFQIEAFGS	KFILDLLNN	GLLSSDYVEI	180
	HYENKPKQYS	KGGEHCYVHG	SIRGVKDSKV	ALSTCNLHGH	MFEDDTFVYM	IEPLELVHDE	240
	KSTGRPHIIQ	KTLAQYYSKQ	MKNLTMERGD	QWPFLLSELQW	LKRRKRAVNP	SRGIFEEMKY	300
	LELMIVNDHK	TYKKHRSSHA	HTNNTAKSVV	NLVDSIYKEQ	LNTRVVLVAV	ETWTEKDQID	360
55	ITTNPVQMLH	EFKYRQRIK	QHADAHVHLIS	RVTFHYKRSS	LSYFGGVCSR	TRGVGVNEYG	420
	LPMAVAQVLS	QSLAQNLGIQ	WEPSSRKPKC	DCETESWGGCI	MEETGVSHSR	KFSKCSILEY	480
	RDFLQRGGGA	CLFNRPTKLF	EPTECGNGYV	EAGEECDGCF	HVECYGLCK	KCSLSNGAHC	540
	SDGPCCNNTS	CLFQPRGYEC	RDVNECDIT	EYCTGDSGQC	PPNLHKQDGY	ACNQNGRCY	600
	NGECKTRDNQ	QYIWTGKAA	GSDKFCYKEL	NTEGTEKGNC	GKDGDRWIQC	SKHDFVFCGFL	660
60	LCTNLTRAPR	IGQLQGEIIP	TSFYHQGRVI	DCSGAHVVLD	DDTDVGYVED	GTPCGPSMMC	720
	LDRKCLQIQA	LNMSSCPLDS	KGKVCSGHGV	CSNEATCICD	FTWAGTDCSI	RDPVRNLHPP	780
	KDEGPKGPSA	TNLIIGSIAG	AILVAAIVLG	GTGWGFKNVK	KRRFDPTQQG	PI	

Seq ID NO: 326 DNA sequence

Nucleic Acid Accession #: AK074418.1

Coding sequence: 244-1515

65	1	11	21	31	41	51	
	CTTTCTCCAA	GACGGCCGGC	CATGCTCTCC	TCCTCTGCCA	GTCTCCTCCA	CCACTCTCTA	60
70	ACCTGAGAGC	CTGTGGAACC	TGCCCGTCTC	CCCTCCTCCA	TCAGACACAC	CTGCCTAGGA	120
	AACAGATGGA	AAAAGTGAGG	GACCGGTGAG	TGACTTGCTG	CTAAAGTTTA	TACCAGATGC	180
	AAATGACAGA	GCTGGAGTTC	TGCTGTGCCT	GGAAAGGACC	TCGGAAGTCT	TCTAAGGAGA	240
	GTCATGGCGT	ATTACCAGGA	GCCTTCAGTG	GAGACCTCCA	TCATCAAGTT	CAAAGACCAG	300
75	GACTTTACCA	CCTTGCGGGA	TCATGCTCTG	AGCATGGGCC	GGACGTTTAA	GGATGAGACA	360
	TTCCCCGAG	CAGATTCTTC	CATAGGCCAG	AAGCTGCTCC	AGGAAAAACG	CCTCTCCAAT	420
	GTGATATGGA	AGCGGCCACA	GATGCTACCA	GGGGGTCTCT	CTCACTTCAT	CCTGGATGAT	480
	ATAAGCAGAT	TTGACATCCA	ACAAGGAGGC	GCAGCTGACT	GCTGGTTCCT	GGCAGCACTG	540
	GGATCCTTGA	CTCAGAACCC	ACAGTACAGG	CAGAAGATCC	TGATGGTCCA	AAGCTTTTCA	600
80	CACCAATATG	CTGGCATTTC	CCGTTTCCGG	TTCTGGCAAT	GTGGCCAGTG	GGTGGAAGTG	660
	GTGATTGATG	ACCGCCTACC	TGTCCAGGGA	GATAAATGCC	TCTTTGTGCG	TCCTCGCCAC	720
	CAAAACCAAG	AGTTCCTGGC	CTGCCTGCTG	GAGAAGGCCT	ATGCCAAGCT	GCTCGGATCC	780
	TATTCGATC	TGCATATGG	CTTCTCGAG	GATGCCCTGG	TGGACCTCAC	AGGAGGCGTG	840
	ATCACCAACA	CTCATCTGCA	CTCTTCCCTT	GTGGACCTGG	TGAAGGCAGT	GAAGACAGCG	900
85	ACCAAGGCAG	GCTCCCTGAT	AACCTGTGCC	ACTCCAAGTG	GGCCAACAGA	TACAGCACAG	960
	GCGATGGAGA	ATGGGCTGGT	GAGTCTCCAT	GCCTACACTG	TGACTGGGGC	TGAGCAGATT	1020
	CAATACCGAA	GGGGCTGGGA	AGAAATTATC	TCCCTGTGGA	ACCCCTGGGG	CTGGGGCGAG	1080
	ACCGAATGGA	GAGGGCGCTG	GAGTGATGGG	TCTCAGGAGT	GGGAGGAAAC	CTGTGATCCG	1140



	CGGAAAAGCC	AGCTACATAA	GAAACGGGAA	GATGGCGAGT	TTTGGATGTC	GTGTCAAGAT	1200
	TTCCAACAGA	AATTCATCGC	CATGTTTATA	TGTAGCGAAA	TTCCAATTAC	CCTGGACCAT	1260
	GGAAACACAC	TCCACGAAGG	ATGGTCCCAA	ATAATGTTTA	GGAAGCAAGT	GATTCTAGGA	1320
5	AACACTGCAG	GAGGACCTCG	GAATGATGCT	CAATTCAACT	TCTCTGTGCA	AGAGCCAATG	1380
	GAAGGCACCA	ATGTTGTCTG	GTGCGTCACA	GTTGCTGTCA	CACCATCAAA	TTTGAAAGCA	1440
	GAAGATGCAA	AATTTCCACT	CGATTTCCAA	GTGATTCTGG	CTGGCTCACA	GAAACACTGT	1500
	CCAAAGCTCA	AATAATAAAT	TCCGCCGCAA	CTTCACCATG	ACTTACCATC	TGAGCCCTGG	1560
	GAAGTATGTT	GTGTTTGCAC	AGACACGGAG	AAAATCAGCG	GAGTTCTTGC	TCCGAATCTT	1620
10	CCTGAAAATG	CCAGACAGTG	ACAGGCACCT	GAGCAGCCAT	TTCAACCTCA	GAATGAAGGG	1680
	AAGCCCTTCA	GAACATGGCT	CCCAACAAAG	CATTTTCAAC	AGATATGCTC	AGCAGGTATG	1740
	GTACCTAGCA	CCCAGGGGCG	TTACGTGGGA	TTGGAGAAAG	GGGACCTGAG	GGAGGGACAG	1800
	CCCTCACAGG	CCCTTACTGG	GATGCAGAGA	GGAGAAGTGA	CTTGATGGAC	TATTTTACCT	1860
	GCCTCTCTTC	CTGGATCGTC	TCCAGAAGTG	CTGTGGCTGC	CAAGCTCGGT	AGAGACGTGG	1920
15	CGCCCCACCC	AGTCTCATCC	GGGGGACTTC	AAGCTGGAAT	GCAGAGCTTA	GAAAGGGAGG	1980
	GGATAATTAT	GGGGTGTGAG	GTGCATTGCC	CTCTAAATCT	TTAAACAAGC	AATTGGCAGT	2040
	ACCCCGTGAA	ACCTTTCTCT	CTCCTACTCG	GCCACCTCCC	ACCAACCTGG	CATCGTTCCT	2100
	CCCGGGAGCT	AGCCAGCTTC	AGAAAGCACA	TACAGCATCC	TTGCTGCCAA	ACCACCTATG	2160
	TGCACACAGG	ATTTCTTTAA	TGGCTTAATA	AACTGTTATA	AAGAACTCCT	TGACTTGTCA	2220
20	GAATAAAATA	GCTGCCAGGG	GCTCTGCACA	ATGAGCCTCT	TACCGTTAAA	AAAAAAAAAA	2280
	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA				

Seq ID NO: 327 Protein sequence:  
Protein Accession #: BAB85075.1

25	1	11	21	31	41	51	
	MAYYQEPSVE	TSIIKFKDQD	FTTLRDHCLS	MGRTFKDETF	PAADSSIGQK	LLQEKRLSNV	60
	IWKRPQDLPG	GPPHFILDDI	SRFDIQQGGA	ADCWFLAALG	SLTQNPQYRQ	KILMVQSFSS	120
	QYAGIFRFRF	WQCGQWVEV	IDDLRPVQGD	KCLFVRPRHQ	NQEFWPCLE	KAYAKLLGSY	180
30	SDLHYGFLED	ALVDLTGGVI	TNIHLHSSPV	DLVKAVKTAT	KAGSLITCAT	PSGPTDTAQA	240
	MENGLVSLHA	YTVTGAEQIQ	YRRGWEEIIS	LWNPWGWGET	EWGRGRWSDS	QEWEEETCDPR	300
	KSQHLKKRED	GEFWMSCQDF	QQKFIAMFIC	SEIPITLDHG	NTLHEGWSQI	MFRKQVILGN	360
	TAGGPRNDAQ	FNFSVQEPME	GTNVVVCVT	AVTPSNLKA	DAKFPLDFQV	ILAGSQKHCP	420
35	KLK						

Seq ID NO: 328 DNA sequence  
Nucleic Acid Accession #: BC017490.1  
Coding sequence: 74-2788

40	1	11	21	31	41	51	
	GTGGGTCACG	TGAACCACTT	TTCGCGCGAA	ACCTGGTTGT	TGCTGTAGTG	GCGGAGAGGA	60
	TGCTGGTACT	GCTATGGCGG	AATCATCGGA	ATCCTTCACC	ATGGCATCCA	GCCCGGCCCA	120
45	GCGTCGGCGA	GGCAATGATC	CTCTACCTC	CAGCCCTGGC	CGAAGCTCCC	GGCGTACTGA	180
	TGCCCTCACC	TCCAGCCCTG	GCCGTGACCT	TCCACCATT	GAGGATGAGT	CCGAGGGGCT	240
	CCTAGGCACA	GAGGGGCCCC	TGGAGGAAGA	AGAGGATGGA	GAGGAGCTCA	TTGGAGATGG	300
	CATGGAAAGG	GACTACCGCG	CCATCCGAGA	GCTGGACGCC	TATGAGGCCG	AGGGACTGGC	360
	TCTGGATGAT	GAGGACGTAG	AGGAGCTGAC	GGCCAGTCAG	AGGGAGGCAG	CAGAGCGGGC	420
50	CATGCGGCAG	CGTGACCGGG	AGGCTGGCCG	GGGCTGGGGC	CGCATGCGCC	GTGGGCTCCT	480
	GTATGACAGC	GATGAGGAGG	ACGAGGAGCG	CCCTGCCCGC	AAGCGCCGCC	AGGTGGAGCG	540
	GGCCACGGAG	GACGCGGAGG	AGGACGAGGA	GATGATCGAG	AGCATCGAGA	ACCTGGAGGA	600
	TCTCAAAGGC	CACTCTGTGC	GCGAGTGGGT	GAGCATGGCG	GGCCCCCGGC	TGGAGATCCA	660
	CCACCGCTTC	AAGAACTTCC	TGCGCACTCA	CGTCGACAGC	CACGGCCACA	ACGTCTTCAA	720
55	GGAGCGCATC	AGCGACATGT	GCAAAGAGAA	CCGTGAGAGC	CTGGTGGTGA	ACTATGAGGA	780
	CTTGGCAGCC	AGGGAGCACG	TGCTGGCCTA	CTTCTGCTCT	GAGGCACCGG	CGGAGCTGCT	840
	GCAGATCTTT	GATGAGGCTG	CCCTGGAGGT	GCTACTGGCC	ATGTACCCCA	AGTACGACCG	900
	CATCACC AAC	CACATCCATG	TCCGCATCTC	CCACCTGCCT	CTGGTGGAGG	AGCTGCGCTC	960
	GCTGAGGCAG	CTGCATCTGA	ACCAGCTGAT	CCGCACCAGT	GGGGTGGTGA	CCAGCTGCAC	1020
60	TGGCGTCCTG	CCCCAGCTCA	GCATGGTCAA	GTACAAGTGC	AACAAGTGCA	ATTTCTGCTC	1080
	GGGTCTCTTC	TGCCAGTCCC	AGAACCAGGA	GGTGAAACCA	GGCTCCTGTC	CTGAGTGCCA	1140
	GTGCGCCGGC	CCCTTTGAGG	TCAACATGGA	GGAGACCATC	TATCAGAACT	ACCAGCGTAT	1200
	CCGAATCCAG	GAGAGTCCAG	GCAAAGTGGC	GGCTGGCCGG	CTGCCCGCT	CCAAGGACGC	1260
	CATTCTCCTC	GCAGATCTGG	TGGACAGCTG	CAAGCCAGGA	GACGAGATAG	AGCTGACTGG	1320
65	CATCTATCAC	AACAACATAT	ATGGCTCCCT	CAACACTGCC	AATGGCTTCC	CTGTCTTTGC	1380
	CACTGTTCAT	CTAGCCAACC	ACGTGGCCAA	GAAGGACAAC	AAGGTTGCTG	TAGGGGAAC	1440
	GACCGATGAA	GATGTGAAGA	TGATCACTAG	CCTCTCCAAG	GATCAGCAGA	TCCGAGAGAA	1500
	GATCTTTGCC	AGCATTTGCT	CTTCCATCTA	TGGTCATGAA	GACATCAAGA	GAGGCCTGGC	1560
	TCTGGCCCTG	TTCCGAGGGG	AGCCCAAAAA	CCCAGGTGGC	AAGCACAAGG	TACGTGGTGA	1620
70	TATCAACGTG	CTCTTGTGCG	GAGACCTGGG	CACAGCGAAG	TCGCAATTTC	TCAAGTATAT	1680
	TGAGAAAGTG	TCCAGCCGAG	CCATCTTCAC	CACTGGCCAG	GGGGCGTCGG	CTGTGGGCCT	1740
	CACGGCGTAT	GTCCAGCGGC	ACCCTGTCAG	CAGGGAGTGG	ACCTTGGAGG	CTGGGGCCCT	1800
	GGTTCTGGCT	GACCGAGGAG	TGTGTCTCAT	TGATGAATTT	GACAAGATGA	ATGACCAGGA	1860
	CAGAACCAGC	ATCCATGAGG	CCATGGAGCA	ACAGAGCATC	TCCATCTCGA	AGGCTGGCAT	1920
75	CGTCACTCC	CTGCAGGCTC	GCTGCAGGAT	CATTGTGCTC	GCCAACCCCA	TAGGAGGGCG	1980
	CTACGACCCC	TCGCTGACTT	TCTCTGAGAA	CGTGACCTC	ACAGAGCCCA	TCATCTCAGC	2040
	CTTTGACATC	CTGTGTGTGG	TGAGGGACAC	CGTGGACCCA	GTCCAGGACG	AGATGCTGGC	2100
	CCGCTTCGTG	GTGGGCAGCC	ACGTACAGAC	CCACCCAGC	AACAAGGAGG	AGGAGGGGCT	2160
	GGCCAATGGC	AGCGCTGCTG	AGCCCGCCAT	GCCCAACACG	TATGGCGTGG	AGCCCTGCC	2220
80	CCAGGAGGTC	CTGAAGAAAGT	ACATCATCTA	CGCCAAGGAG	AGGGTCCACC	CGAAGCTCAA	2280
	CCAGATGGAC	CAGGACAAGG	TGGCCAAGAT	GTACAGTGAC	CTGAGGAAAG	AATCTATGGC	2340
	GACAGGCAGC	ATCCCCATTA	CGGTGCGGCA	CATCGAGTCC	ATGATCCGCA	TGGCGGAGGC	2400
	CCACGCGCGC	ATCCATCTGC	GGGCTATGAT	GATCGAAGAC	GACGTCAACA	TGGCCATCCG	2460
	CGTGATGCTG	GAGAGCTTCA	TAGACACACA	GAAGTTTCA	GTCATGCGCA	GCATGCGCAA	2520
85	GACTTTTGCC	CGCTACCTTT	CATTCCGGCG	TGACAACAAT	GAGCTGTTGC	TCTTCATACT	2580
	GAAGCAGTTA	GTGGCAGAGC	AGGTGACATA	TCAGCGCAAC	CGCTTTGGGG	CCCAGCAGGA	2640
	CACTATTGAG	GTCCCTGAGA	AGGACTTGGT	GGATAAGGCT	CGTCAGATCA	ACATCCACAA	2700



CCTCTCTGCA TTTTATGACA GTGAGCTCTT CAGGATGAAC AAGTTCAGCC ACGACCTGAA 2760  
 AAGGAAAATG ATCCTGCAGC AGTTCTGAGG CCCTATGCCA TCCATAAGGA TTCCTTGGGA 2820  
 TTCTGGTTTG GGGTGGTCAG TGCCCTCTGT GCTTTATGGA CACAAAACCA GAGCACTTGA 2880  
 TGAACTCGGG GTACTAGGGT CAGGGCTTAT AGCAGGATGT CTGGCTGCAC CTGGCATGAC 2940  
 5 TGTGTTGTTT TCCAAGCCTG CTTTGTGCTT CTCACCTTTG GGTGGGATGC CTTGCCAGTG 3000  
 TGTCTTACTT GGTGCTGAA CATCTTGCCA CCTCCGAGTG CTTTGTCTCC ACTCAGTACC 3060  
 TTGGATCAGA GCTGCTGAGT TCAGGATGCC TGCGTGTGGT TTAGGTGTTA GCCTTCTTAC 3120  
 ATGGATGTCA GGAGAGCTGC TGCCCTCTTG GCGTGAGTTG CGTATTCAGG CTGCTTTTGC 3180  
 10 TGCCTTTGGC CAGAGAGCTG GTTGAAGATG TTTGTAATCG TTTTCAGTCT CCTGCAGGTT 3240  
 TCTGTGCCCC TGTGGTGGAA GAGGGCACGA CAGTGCCAGC GCAGCGTTCT GGGCTCCTCA 3300  
 GTCGCAGGGG TGGGATGTGA GTCATGCGGA TTATCCACTC GCCACAGTTA TCAGCTGCCA 3360  
 TTGCTCCCTG TCTGTTTCCC CACTCTCTTA TTTGTGCATT CGGTTTGGTT TCTGTAGTTT 3420  
 TAATTTTAA TAAAGTTGAA TAAATATAA AAAAAAAAAA AAAAAA

Seq ID NO: 329 Protein sequence:  
 Protein Accession #: AAH17490.1

1 11 21 31 41 51  
 | | | | |  
 20 MAESSESFTM ASSPAQRRRG NDPLTSSPGR SSRRTDALTS SPGRDLPPFE DESEGLLGTE 60  
 GPLEEEEDGE ELIGDGMERD YRAIPELDAY EAEGALDDE DVEELTASQR EAAERAMRQR 120  
 DREAGRGLGR MRRGLLYDS EDEEREPARK RRQVERATED GEEDEEMIES IENLEDLKGH 180  
 SVREWVSMAG PRLEIHHRFK NFLRTHVDSH GHNVFKERIS DMCKENRESL VVNYEDLAAR 240  
 25 EHVLAFLPE APAELLQIFD EAALEVVLAM YPKYDRITNH IHVRISHLPL VEELRSLRQL 300  
 HLNQLIRTSV VVTSCITGVL QLSMVKYN CN KCFVLGPFQ QSQNQEVKPG SCPECQSAGP 360  
 FEVNMEETIY QNYQRIRIQE SPQKVAAGRL PRSKDAILLA DLVDSCKPGD EIELTGIYHN 420  
 NYDGSNLNAN GFPVFATVIL ANHVAKKDNK VAVGELTDED VKMITSLSKD QQIGEKIFAS 480  
 IAPSIYGHED IKRGLALALF GGEKPNPGGK HKVRGDINVL LCGDPGTAKS QFLKYIEKVS 540  
 30 SRAIFTTGGG ASAVGLTAYV QRHPVSREW T LEAGALVLAD RGVCLIDEFD KMNDQDRTSI 600  
 HEAMEQQSIS ISKAGIVTSL QARCTVIAAA NPIGGRYDPS LTFSENVDLT EPIISRFDIL 660  
 CVVRDITVDPV QDEMLARFV GSHVRHHPN KEEGLANGS AAEPAMPNTY GVEPLPQEV 720  
 KKYIYAKER VHPKLNQMDQ DKVAKMYS DL RKESMATGSI PITVRHIESM IRMAEAHARI 780  
 HLRDYVIEDD VNMAIRVMLE SFIDTQKFSV MRSMRKT FAR YLSFRDNNE LLLFILKQLV 840  
 35 AEQVTYQRNR FGAQQDTIEV PEKDLVDKAR QINIHNLSAF YDSELFMRNK FSHDLKRKMI 900  
 LQQF

Seq ID NO: 330 DNA sequence  
 Nucleic Acid Accession #: M17254  
 Coding sequence: 257-1645

1 11 21 31 41 51  
 | | | | |  
 45 GTCCGCGCGT GTCCGCGCCC GCGTGTGCCA GCGCGCGTGC CTTGGCCGTG CGCGCCGAGC 60  
 CCGGTCGCAC TAACTCCCTC GCGCGCCAGC GCGCGCGTAA CCTCTCGGTT ATTCCAGGAT 120  
 CTTTGGAGAG CCGAGGAAAG CCGTGTGAC CAAAAGCAAG ACAATGACT CACAGAGAAA 180  
 AAAGATGGCA GAACCAAGGG CAACTAAAGC CGTCAGGTTT TGAACAGCTG GTAGATGGGC 240  
 TGGCTTACTG AAGGACATGA TTCAGACTGT CCCGGACCCA GCAGCTCATA TCAAGGAAGC 300  
 CTTATCAGTT GTGAGTGAGG ACCAGTCTGT GTTTGAGTGT GCCTACGGAA CGCCACACCT 360  
 50 GGCTAAGACA GAGATGACCG CGTCCCTCCTC CAGCGACTAT GGACAGACTT CCAAGATGAG 420  
 CCCACGCGTC CCTCAGCAGG ATTGGCTGTC TCAACCCCTC GCCAGGGTCA CCATCAAAAT 480  
 GGAATGTAAC CCTAGCCAGG TGAATGGCTC AAGGAACCTC CCTGATGAAT GCATGTGGC 540  
 CAAAGGCGGG AAGATGGTGG GCAGCCGAGA CACCGTTGGG ATGAACTACG GCAGCTACAT 600  
 GGAGGAGAAG CACATGCCAC CCCCACCAT GACCACGAAC GAGCGCAGAG TTATCGTGCC 660  
 55 AGCAGATCCT ACGCTATGGA GTACAGACCA TGTGCGGCAG TGGCTGGAGT GGGCGGTGAA 720  
 AGAATATGGC CTTCCAGACG TCAACATCTT GTTATTCCAG AACATCGATG GGAAGGAAC 780  
 GTGCAAGATG ACCAAGGACG ACTTCCAGAG GCTCACCCCT AGCTACAACG CCGACATCCT 840  
 TCTCTCACAT TCCACTACC TCAGAGAGAC TCCTCTTCCA CATTTGACTT CAGATGATGT 900  
 TGATAAAGCC TTACAAAAC CTCCACGGTT AATGCTGCT AGAAACACAG ATTTACCATA 960  
 60 TGAGCCCCC AGGAGATCAG CCTGGACCGG TCACGGCCAC CCCACGCCCC AGTCGAAAGC 1020  
 TGCTCAACCA TCTCCTTCCA CAGTGCCCAA AACTGAAGAC CAGCGTCTCT AGTTAGATCC 1080  
 TTATCAGATT CTTGGACCAA CAAGTAGCCG CCTTGCAAAT CCAGGCAGTG GCCAGATCCA 1140  
 GCTTTGGCAG TTCCTCCTGG AGCTCCTGTC GGACAGCTCC AACTCCAGCT GCATCACCTG 1200  
 GGAAGGCACC AACGGGGAGT TCAAGATGAC GGATCCCGAC GAGGTGGGCC GCGCTGGGG 1260  
 65 AGAGCGGAAG AGCAAAACCA ACATGAAGCT AGCCGCGCCC TCCGTTACTA 1320  
 CTATGACAAG AACATCATGA CCAAGGTCCA TGGGAAGCGC TACGCCTACA AGTTGACTT 1380  
 CCACGGGATC GCCCAGGCC TCCAGCCCCA CCCCCCGGAG TCATCTCTGT ACAAGTACCC 1440  
 CTCAGACCTC CCGTACATGG GCTCCTATCA CGCCACCCA CAGAAGATGA ACTTTGTGGC 1500  
 GCCCCACCTT CCAGCCCTCC CCGTGACATC TTCCAGTTTT TTTGCTGCCC CAAACCCATA 1560  
 CTGGAATTCA CCAACTGGGG GTATATAACC CAACACTAGG CTCCCACCA GCCATATGCC 1620  
 70 TTCTCATCTG GCACTTACT ACTAAAGACC TGCGGAGGCT TTTTCCATC AGCGTGCAAT 1680  
 CACCAGCCCA TCGCCACAAA CTCTATCGGA GAACATGAAT CAAAAGTGCC TCAAGAGGAA 1740  
 TGAAAAAAGC TTTACTGGGG CTGGGGAAGG AAGCCGGGGA AGAGATCCAA AGACTCTTGG 1800  
 GAGGGAGTTA CTGAAGTCTT ACTACAGAAA TGAGGAGGAT GCTAAAAATG TCACGAATAT 1860  
 75 GGACATATCA TCTGTGGACT GACCTTGTA AAGACAGTGT ATGTAGAAGC ATGAAGTCTT 1920  
 AAGGACAAAG TGCCAAAGAA AGTGGTCTTA AGAAATGTAT AAACCTTAGA GTAGAGTTTG 1980  
 AATCCCACTA ATGCAAACTG GGATGAACTA AAAGCAATAG AAACAACACA GTTTTGACCT 2040  
 AACATAACGT TTATAATGCC ATTTTAAAGG AAACACTCTG TATTTAAAA TAGTTTCATA 2100  
 TCAAAAACAA GAGAAAAGAC ACGAGAGAGA CTGTGGCCCA TCAACAGACG TTGATATGCA 2160  
 80 ACTGCATGGC ATGTGCTGTT TTGGTTGAAA TCAAAATACAT TCCGTTTGAT GGACAGCTGT 2220  
 CAGCTTTCTC AAAGTGTGAA GATGACCCAA AGTTTCCAAC TCCTTTACAG TATTACCGGG 2280  
 ACTATGAAGT AAAAGGTGGG ACTGAGGATG TGTATAGAGT GAGCGTGTGA TTGTAGACAG 2340  
 AGGGGTGAAG AAGGAGGAGG AAGAGGCAGA ACCAGGCTGG GAAAGAAACT 2400  
 TCTCAAGCAA TGAAGACTGG ACTCAGGACA TTTGGGGACT GTGTACAATG AGTTATGGAG 2460  
 85 ACTCGAGGGT TCATGCAGTC AGTGTATAT CAAACCCAGT GTTAGGAGAA AGGACACAGC 2520  
 GTAATGGAGA AAGGGAAGTA GTAGAAATCA GAAACAAAAA TGCGCATCTC TTTCTTTGTT 2580  
 TGTCAAATGA AAATTTTAAAC TGGAATGTGC TGATATTTAA GAGAAACATT CAGGACCTCA 2640  
 TCATTATGTG GGGGCTTTGT TCTCCACAGG GTCAGGTAAG AGATGGCCTT CTTGGCTGCC 2700



ACAATCAGAA ATCACGCAGG CATTTTGGGT AGGCGGCCCTC CAGTTTTCCT TTGAGTCGCG 2760  
 AACGCTGTGC GTTTGTTCAGA ATGAAGTATA CAAGTCAATG TTTTTCCTCC TTTTATATA 2820  
 ATAATTATAT AACTTATGCA TTTATACACT ACAGATTGAT CTCGGCCAGC CAAAGACACA 2880  
 CGACAAAAGA GACAATCGAT ATAATGTGGC CTTGAATTTT AACTCTGTAT GCTTAATGTT 2940  
 TACAATATGA AGTTATTAGT TCTTAGAATG CAGAATGTAT GTAATAAAAT AAGCTTGGCC 3000  
 TAGCATGGCA AATCAGATTT ATACAGGAGT CTGCATTTGC ACTTTTTTTA GTGACTAAAG 3060  
 TTGCTTAATG AAAACATGTG CTGAATGTTG TGGATTTTGT GTTATAATTT ACTTTGTCCA 3120  
 GGAACCTGTG CAAGGGAGAG CCAAGGAAAT AGGATGTTTG GCACCC

Seq ID NO: 331 Protein sequence  
 Protein Accession #: AAA52398

1 11 21 31 41 51  
 MIQTVDPDAA HIKEALSUVS EDQSLFECAY GTPHLAKTEM TASSSSDYGQ TSKMSPRVPQ 60  
 QDWLSQPPAR VTIKMECNPS QVNGSRNSPD ECSVAKGGKM VGSPDTVGMN YGSYMEEKHM 120  
 PPPNMTTNER RVIVPADPTL WSTDHVRQWL EWAVKEYGLP DVNILLFQNI DGKELCKMTK 180  
 DDFQRLTPSY NADILLSHLH YLRETPLPHL TSDDVDKALQ NSPRLMHARN TDLPEYPPRR 240  
 SAWTGHGHPT PQSKAAQPSF STVPKTEDQR PQLDPYQILG PTSSRLANPG SGQIQLWQFL 300  
 LELLSDDSSNS SCITWEGTNG EFKMTDPDEV ARRWGERKSK PNMNYDKLSR ALRYYDKNI 360  
 MTKVHGKRYA YKFDPHGIAQ ALQPHPESS LYKYPSDLPY MGSYHAHPQK MNFVAPHPPA 420  
 LPVTSSSFFA APNPYWNSTP GGIYPNTRLP TSHMPSHLGT YY 462

Seq ID NO: 332 DNA sequence  
 Nucleic Acid Accession #: NM\_000020  
 Coding sequence: 283-1794

1 11 21 31 41 51  
 AGGAAACGGT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA 60  
 AGAAACATTT TTGCTCCAGC CCCCATCCCA GTCCCGGGAG GCTGCCGCGC CAGCTGCGCC 120  
 GAGCGAGCCC CTCCCCGGCT CCAGCCCAGT CCGGGGCCGC GCCGGACCCC AGCCCGCCGT 180  
 CCAGCGCTGG CGGTGCAACT GCGGCCGCGC GGTGGAGGGG AGGTGGCCCC GGTCCGCCGA 240  
 AGGCTAGCGC CCGGCCACCC GCAGAGCGGG CCAGAGGGA CCATGACCTT GGGCTCCCCC 300  
 AGGAAAGGCC TTCTGATGCT GCTGATGGCC TTGGTGACCC AGGGAGACCC TGTGAAGCCG 360  
 TCTCGGGGCC CGCTGGTGAC CTGCACGTGT GAGAGCCAC ATTGCAAGGG GCCTACCTGC 420  
 CGGGGGCCCT GGTGCACAGT AGTGCTGGTG CCGGAGGAGG GGAGGCACCC CCAGGAACAT 480  
 CGGGGCTGCG GGAACCTGCA CAGGGAGCTC TGCAGGGGGC GCCCCACCGA GTTCGTCAAC 540  
 CACTACTGCT GCGACAGCCA CCTCTGCAAC CACAACGTGT CCCTGGTGCT GGAGGCCACC 600  
 CAACCTCCTT CCGAGCAGCC GGAACAGAT GGCCAGCTGG CCCTGATCCT GGGCCCCGTG 660  
 CTGGCCTTGC TGGCCCTGGT GGCCCTGGGT GTCTGGGCC TGTGGCATGT CCGACGGAGG 720  
 CAGGAGAAGC AGCGTGGCCT GCACAGCGAG CTGGGAGAGT CCAGTCTCAT CCTGAAAGCA 780  
 TCTGAGCAGG GCGACACGAT GTTGGGGGAC CTCTGGACA GTGACTGCAC CACAGGGAGT 840  
 GGCTCAGGGC TCCCCTTCTT GGTGCAGAGG ACAGTGGCAC GGCAGGTGTC CTTGGTGGAG 900  
 TGTGTGGGAA AAGGCCGCTA TGGCGAAGTG TGGCGGGGCT TGTGGCACGG TGAGAGTGTG 960  
 GCCGTCAAGA TCTTCTCCTC GAGGGATGAA CAGTCTGGT TCCGGGAGAC TGAGATCTAT 1020  
 AACACAGTAT TGCTCAGACA CGACAACATC CTAGGCTTCA TCGCCTCAGA CATGACCTCC 1080  
 CGCAACTCGA GCACGCAGCT GTGGCTCATC ACGCACTACC ACGAGCACGG CTCCCTCTAC 1140  
 GACTTTCTGC AGAGACAGAC GCTGGAGCCC CATCTGGCTC TGAGGCTAGC TGTGTCCGCG 1200  
 GCATGCGGCC TGGCGCACCT GCACGTGGAG ATCTTCGGTA CACAGGGCAA ACCAGCCATT 1260  
 GCCCACCGCG ACTTCAAGAG CCGCAATGTG CTGGTCAAGA GCAACCTGCA GTGTTGCATC 1320  
 GCCGACCTGG GCCTGGCTGT GATGCACTCA CAGGGCAGCG ATTACCTGGA CATCGGCAAC 1380  
 AACCAGAGAG TGGGCACCAA GCGGTACATG GCACCCGAGG TGCTGGACGA GCAGATCCGC 1440  
 ACGGACTGCT TTGAGTCCTA CAAGTGGACT GACATCTGGG CCTTTGGCCT GGTGCTGTGG 1500  
 GAGATTGCCC GCCGGACCAT CGTGAATGGC ATCGTGGAG ACTATAGACC ACCCTTCTAT 1560  
 GATGTGGTGC CCAATGACCC CAGCTTTGAG GACATGAAGA AGGTGGTGTG TGTGGATCAG 1620  
 CAGACCCCA CCATCCCTAA CCGGCTGGCT GCAGACCCGG TCCTCTCAGG CCTAGCTCAG 1680  
 ATGATGCGGG AGTGCTGGTA CCCAAACCCC TCTGCCCGAC TCACCGCGCT GCGGATCAAG 1740  
 AAGACACTAC AAAAAATTAG CAACAGTCCA GAGAAGCCTA AAGTGATTCA ATAGCCAGG 1800  
 AGCACCTGAT TCCTTTCTGC CTGCAGGGGG CTGGGGGGGT GGGGGGCAGT GGATGGTGCC 1860  
 CTATCTGGGT AGAGGTAGTG TGAGTGTGGT GTGTGCTGGG GATGGGCAGC TGCCTCTGCC 1920  
 TGCTCGGCCC CCAGCCCACC CAGCCAAAAA TACAGCTGGG CTGAAACCTG ATCCCTTGCT 1980  
 GTCTGGCCTG CTCAAAGCGG CAGGCTCCCT GACGCTGGC TCTCTCCCA CCCCTATGGC 2040  
 CAGCATGGTG CACCCCTAC ACTCCCAGG ACAGGATGCA AAAGAGGCTC CAGAGTCAGA 2100  
 GTGCCAAGCC AGGGAATCCC AGTCCCAGC TCAGAGCCCG GGCCTGCACT TTGCCCTG 2160  
 CCCTTGATCA ACCCACTGC CCCACCAGAG CTGCCAGGGT GGCACAGGGC CCTGTCCAGC 2220  
 CCCTGGCACA CACTTCCCTG CCAGGCCTCA GCCTTAGCA TAAGCTCCAG AGAGCCAGGG 2280  
 CCCATCAGTT TCTCTCTGTG GATTTGTATC TCAGCTCCAT GATGCCTTGG GCTTCTGTG 2340  
 TCCTCAACAA GAGTGCAGCT TGCTGAATGT CAGCTGCCTG AGAGAGCTGG GGCCTGACTT 2400  
 ACTAGGCGAT TAAATCCTAA GAGGTCCTAC TGAGGTGTGG CAGGATCACA GGCCAGTGGA 2460  
 AAAAGGCGAG GTCAGATGGG CAAGGCCAG GACTTTCAGA TTAAGTGA GAATATCGAG 2520  
 GCCAAGCATG GCAGGGGGAA GGTCAAGTGG TGTCAAGAGA CCCAGGTCTG ACCCCGATG 2580  
 TTTGCTCCAT GTGACAAAAG CAGGCCTGTC TCAGGACCTT TTCTTTTCTT TTTTCTTCT 2640  
 TTTTTTTTTT GACACGGAGT TTCGCTCTTG TTGTCCAGGC TAGAGTGCAA TGGCATGATC 2700  
 CCAGCTCACC GCAACGTCTA CCTCCCAGGT TCAAATCATT CTCTTGCTCT AGACTCCCGA 2760  
 GTAGCTGGGA TTACAGGCAC ATGCCACCAT GCCTGGCTAA TTTTGTATAT TTAGTAGAAA 2820  
 CAGGGTTTCA CCATGCTGGC CATGCTGGTT CTCGAACTCC TGACCTCAGG TGTTCACCT 2880  
 ACCTCAGCCT CCCAAAGTGC TGGGGTTACA GGTGTGAGCC ATCGCGCCTG GCCAGGACCT 2940  
 TTGTTTCTTA TCTACATATT GGAAGATTGG GTCTTGATGT CCTTTGAGGC TTCTTTAGCT 3000  
 CTAGTTCTCT GACACTTCAG CCTATATCAC AGCTAACTTC YTCAGTCTCA TCTATTCCTT 3060  
 ATGCTCCAGC CCCTGGCAAT TTGCCTCAAG ATGGGGGTTT GAAAATAACT TTACCTGACT 3120  
 CAAGGAGTGT CTGGAGCACC TCCTAGTCTA AGTCTGCAAG CTCCAGTTCT TGCCATAAAC 3180  
 CATGCCAGTG GCCACCCTTG GGCTCAGACA GCTCTGGGCC TTTTGACCAC AAGCCAGCCC 3240  
 CTCGCCCTCT CTGTGGCATA GTCTTCTCTG CCCCAGGACT GCAGGGCGGC TTCTTCCAAG 3300  
 GCTTCCAAGG CTCAAAAGAA ATTTGGCTCC ATCCAAGAAG GCTCCAGCTC CCCTACTGGC 3360  
 CCCTGGCTTC AGGCCACAC CCCTGGGCCA GGSCCAGAGA GTGTGTCTCA GGAGAATTCA 3420  
 ATGGGCTCTA GAGAGACACA CAGAAAGTTT GGGCATTTGG GAAATTTTCA AGGRTGTATG 3480



TATGGYTAC GTATGGWGCA GGTGTGCTCTG GTCCYKGGGT GCAGGGAAGT GGGCTGCAGG 3540  
 GAAGTGGATT GGAGGGGAGC TTGAGGAATA TAAGGAGCGG GGGTGGAGAC TCAGGCTATG 3600  
 GACAAGGACA GGGGCAAGGT TGGGAAGACC TGGCCTTAGT CGTCTCAGC CTAGGGGAGG 3660  
 GCAGTGAAGA AAGCTCTCCC CGCTCTGCT GTAATGACCC AGAGTAGCCT CCCAGGCCG 3720  
 GCATCTTATG TGTGTCTTCC ACCATCTCTCA TGGTGGCACT TTTCTAGGCC TGTCTCCCAG 3780  
 CATTGTGCAA GGCTCGGAAG AGAACCAGGA AGTGAAACTG GGTGAAAACA GAAAGCTCAA 3840  
 TGGATGGGCT AGGTTCACAG ATCATTAGGG CAGAGTTTGC ACGTCCTCTG GTTCACTGGG 3900  
 AATCCACCCA GGGCAGCAAT CATCTCCCTC TTTGAAGGAT TTTWATTTCT ACTGGGTTTT 3960  
 GGAACAAACT CCTGCTGAGA CCCACAGCC AGAAACTGAA AGCAGCAGCT CCCCAAAGCC 4020  
 TGGAAAATCC CTAAGAGAAG GCCTGGGGGA MAGGAAKTGG AGTGACAGGG GACAGGTAGA 4080  
 GAGAAGGGGG CCCAATGGCC AGGGAGTGAA GGAGGTGGCG TTGCTGAGAG CAGTCTGCAC 4140  
 ATGCTTCTGT CTGAGTGCAG GAAGGTGTTC CAGGTCGAA ATTACACTTC TCGTACCTGG 4200  
 AGACGCTGTT TGTGGGAGCA CTGGGCTCAT GCCTGGCACA CAATAGGTCT GCAATAAACC 4260  
 ATGGTTAAAT CCTGAAAAA AAAAAAAA

Seq ID NO: 333 Protein sequence  
 Protein Accession #: NP\_000011

1 11 21 31 41 51  
 | | | | |  
 MTLGSPRKGL LMLLMALVTQ GDPVKPSRGP LVTCTCESPH CKGPTCRGAW CTVVVLVREEG 60  
 RHPQEHRCG NLHRELRCGR PTEFVNHYCC DSHLCNHVS LVLEATQPPS EQPGTDGQLA 120  
 LILGPVLALL ALVALGVGL WHVRRRQEKQ RGLHSELGES SLILKASEQG DTMLGDLLDS 180  
 DCTTSGSGSL PFLVQRTVAR QVALVECVGK GRYGEVWRGL WHGESVAVKI FSSRDEQSWF 240  
 RETEIYNTVL LRHDNILGFI ASDMTSRNSS TQLWLITHYH EHGSLYDFLQ RQTLLEPHLAL 300  
 RLAVSAACGL AHLHVEIFGT QGKPAIAHRD FKSRNVLVKS NLQCCIADLG LAVMHSQGS 360  
 YLDIGNNPRV GTKRYMAPEV LDEQIRTDCE ESYKWTDIWA FGLVLWEIAR RTIVNGIVED 420  
 YRPPFYDVVP NDPSEFDMKK VVCVDQQTPT IPNRLAADPV LSGLAQMMRE CWYPNPSARL 480  
 TALRIKKTLO KISNSPEKPK VIQ

Seq ID NO: 334 DNA sequence  
 Nucleic Acid Accession #: NM\_004126.1  
 Coding sequence: 108-329

1 11 21 31 41 51  
 | | | | |  
 GGCACGAGCT CGTGCCGGCC TTCAGTTGTT TCGGGACGCG CCGAGCTTCG CCGCTCTTCC 60  
 AGCGGCTCCG CTGCCAGAGC TAGCCCAGAG CCGGTTCTGG GCGGAAAATG CCTGCCCTTC 120  
 ACATCGAAGA TTTGCCAGAG AAGGAAAAAC TGAAATGGA AGTTGAGCAG CTTGCGAAAG 180  
 AAGTGAAGTT GCAGAGACAA CAAGTGTCTA AATGTTCTGA AGAAATAAAG AACTATATTG 240  
 AAGAACGTTT TGGAGAGGAT CCTCTAGTAA AGGGAATTCC AGAAGACAAG AACCCTTTTA 300  
 AAGAAAAGG CAGCTGTGTT ATTTCAATAA TAAGTTGGGA GAACTGCAT CCTAAGTGGA 360  
 AGAACTAGTT TGTTTTAGTT TCCCAGATA AAACCAACAT GCTTTTAAAG GAAGGAAGAA 420  
 TGAAATTAAG AGGAGACTTT CTTAAGCACC ATATAGATAG GGTATGTAT AAAAGCATAT 480  
 GTGCTACTCA TCTTTGCTCA CTATGCAGTC TTTTAAAGA GAGCAGAGAG TATCAGATGT 540  
 ACAATTATGG AATAAGAAC ATTACTTGAG CATGACACTT CTTTCAGTAT ATTGCTTGAT 600  
 GCTTCAAATA AAGTTTGTGC TT

Seq ID NO: 335 Protein sequence  
 Protein Accession #: NP\_004117.1

1 11 21 31 41 51  
 | | | | |  
 MPALHIEDLP EKEKLKMEVE QLRKEVKLQR QOVSKCSEEI KNYIEERSGE DPLVKGIPED 60  
 KNPFKEKGGC VIS

Seq ID NO: 336 DNA sequence  
 Nucleic Acid Accession #: NM\_005795  
 Coding sequence: 555-1940

1 11 21 31 41 51  
 | | | | |  
 GCACGAGGGA ACAACCTCTC TCTCTSCAGC AGAGAGTGTC ACCCTCTGCT TTAGGACCAT 60  
 CAAGCTCTGC TAACTGAATC TCATCCTAAT TGCAGGATCA CATTGCAAAG CTTTCACTCT 120  
 TTCCACCTT GCTTGTGGGT AAATCTCTTC TGCGGAATCT CAGAAAGTAA AGTTCCATCC 180  
 TGAGAATATT TCACAAAGAA TTTCCTTAAG AGCTGGACTG GGTCTTGACC CCTGGAATTT 240  
 AAGAAATCT TAAAGACAAT GTCAAATATG ATCCAAGAGA AAATGTGATT TGAGTCTGGA 300  
 GACAATTGTG CATATCGTCT AATAATAAAA ACCCATACTA GCCTATAGAA AACAATATTT 360  
 GAATAATAAA AACCCATACT AGCCTATAGA AAACAATATT TGAAAGATTG CTACCACTAA 420  
 AAAGAAAAC ACTACAACCT GACAAGACTG CTGCAAACTT CAATTGGTCA CCACAACCTG 480  
 ACAAGGTTGC TATAAAACAA GATTGCTACA ACTTCTAGTT TATGTTATAC AGCATATTTT 540  
 ATTTGGGCTT AATGATGGAG AAAAAGTGTA CCCTGTATTT TCTGGTTCTC TTGCCTTTTT 600  
 TTATGATTCT TGTTACAGCA GAATTAGAAG AGAGTCCTGA GGAATCAATT CAGTTGGGAG 660  
 TTACTAGAAA TAAATCATG ACAGCTCAAT ATGAATGTTA CCAAAGATT ATGCAAGACC 720  
 CCATTCAACA AGCAGAAGGC GTTTACTGCA ACAGAACCTG GGATGGATGG CTCTGCTGGA 780  
 ACGATGTTGC AGCAGGAAC GAATCAATGC AGCTGTGCC TGATTACTTT CAGGACTTTG 840  
 ATCCATCAGA AAAAGTTACA AAGATCTGTG ACCAAGATGG AAAGTGGTTT AGACATCCAG 900  
 CAAGCAACAG AACATGGACA AATTATACCC AGTGTAATGT TAACACCCAC GAGAAAGTGA 960  
 AGACTGCACT AAATTTGTTT TACCTGACCA TAATTGGACA CGGATTGTCT ATTGCATCAC 1020  
 TGCTTATCTC GCTTGGCATA TTCTTTTATT TCAAGAGCCT AAGTTGCCAA AGGATTACCT 1080  
 TACACAAAAA TCTGTTCTTC TCATTTGTTT GTAACCTCTG TGTAAACAATC ATTCACCTCA 1140  
 CTGCAGTGGC CAACAACAG GCCTTAGTAG CCACAAATCC TGTTAGTTGC AAAGTGTCCC 1200  
 AGTTCATTCA TCTTTACCTG ATGGGCTGTA ATTACTTTTG GATGCTCTGT GAAGGCATTT 1260  
 ACCTACACAC ACTCATTGTG GTGGCCGTGT TTGCAGAGAA GCAACATTTA ATGTGGTATT 1320  
 ATTTCTTTGG CTGGGGATTT CCACTGATTC CTGCTGTAT ACATGCCATT GCTAGAAGCT 1380



5  
10  
15  
20  
25

TATATTACAA	TGACAATTGC	TGGATCAGTT	CTGATACCCA	TCTCCTCTAC	ATTATCCATG	1440
GCCCAATTTG	TGCTGCCTTTA	CTGGTGAATC	TTTTTTTCTT	GTTAAATATT	GTACGCGTTC	1500
TCATACCAAA	GTTAAAAGTT	ACACACCAAG	CGGAATCCAA	TCTGTACATG	AAAGCTGTGA	1560
GAGCTACTCT	TATCTTGGTG	CCATTGCTTG	GCATTGAATT	TGTGCTGATT	CCATGGCGAC	1620
CTGAAGGAAA	GATTGCAGAG	GAGGTATATG	ACTACATCAT	GCACATCCTT	ATGCACTTCC	1680
AGGGTCTTTT	GGTCTCTACC	ATTTTCTGCT	TCTTTAATGG	AGAGGTTCAA	GCAATTCTGA	1740
GAAGAACTG	GAATCAATAC	AAAATCCAAT	TTGGAAACAG	CTTTTCCAAC	TCAGAAGCTC	1800
TTCTAGTGC	GTCTTACACA	GTGTCAACAA	TCAGTGTATG	TCCAGGTTAT	AGTCATGACT	1860
GTCTAGTGA	ACACTTAAAT	GGAAAAAGCA	TCCATGATAT	TGAAAATGTT	CTCTTAAAC	1920
CAGAAAATTT	ATATAATTGA	AAATAGAAGG	ATGGTTGTCT	CACTGTTTGG	TGCTTCTCCT	1980
AACTCAAGGA	CTTGGACCCA	TGACTCTGTA	GCCAGAAGAC	TTCAATATTA	AATGACTTTG	2040
GGGAATGTCA	TAAAGAAGAG	CCTTCACATG	AAATTAGTAG	TGTGTTGATA	AGAGTGTAAC	2100
ATCCAGCTCT	ATGTGGGAAA	AAAGAAATCC	TGGTTTGTA	TGTTTGTGAG	TAAATACTCC	2160
CACTATGCCT	GATGTGACGC	TACTAACCTG	ACATCACCAA	GTGTGGAATT	GGAGAAAAGC	2220
ACAATCAACT	TTTCTGAGCT	GGTGTAGGCC	AGTTCCAGCA	CACCATTGAT	GAATTCAAAC	2280
AAATGGCTGT	AAAATAAAC	ATACATGTTG	GGCATGATTC	TACCTTATT	CSCCCCAAGA	2340
GACCTAGCTA	AGGTCTATAA	ACATGAAGGG	AAAATTAGCT	TTTAGTTTTA	AAACTCTTTA	2400
TCCCATCTTG	ATTGGGGCAG	TTGACTTTTT	TTTTTTCCCA	GAGTGCCGTA	GTCTTTTTTG	2460
TAACTACCTT	CTCAAATGGA	CAATACCAGA	AGTGAATTAT	CCCTGCTGGC	TTTCTTTTCT	2520
CTATGAAAAG	CAACTGAGTA	CAATTGTTAT	GATCTACTCA	TTTGCTGACA	CATCAGTTAT	2580
ATCTTGTGGC	ATATCCATTG	TGGAACCTGG	ATGCAACAGG	TGTATAATAT	GCAATCTTAC	2640
TTCTATATCA	TTAGGAAAAC	ATCTTAGTTG	ATGCTACAAA	ACACCTTGTC	AACCTCTTCC	2700
TGTCTTACCA	AACAGTGGGA	GGGAATTCCT	AGCTGTAAAT	ATAAATTTTG	CCCTTCCATT	2760
TCTACTGTAT	AAACAAATTA	GCAATCATTT	TATATAAAGA	AAATCAATGA	AGGATTTCTT	2820
ATTTTCTTGG	AATTTTGTAA	AAAGAAATTG	TGAAAAATGA	GCTTGTAAT	ACTCCATTAT	2880
TTTATTTTAT	AGTCTCAAAT	CAAATACATA	CAACCTATGT	AATTTTAAA	GCAAATATAT	2940
AATGCAACAA	TGTGTGTATG	TTAATATCTG	ATACTGTATC	TGGGCTGATT	TTTTAAATAA	3000
AATAGAGTCT	GGAATGCT					

Seq ID NO: 337 protein sequence  
Protein Accession #: NP\_005786.1

35  
40

1	11	21	31	41	51	
MEKKCTLYFL	VLLPFFMILV	TAELEESPED	SIQLGVTRNK	IMTAQYECYQ	KIMQDPIQQA	60
EGVYCNRTWD	GWLCWNDVAA	GTESMQLCPD	YFQDFDPSEK	VTKICDQDGN	WFRHPASNRT	120
WTNYTQCNVN	THEKVKTALN	LFYLTIIHGH	LSIASLLISL	GIFFYFKSL	CQRITLHKNL	180
FFSFVCNSVV	TIIHLTAVAN	NQALVATNPV	SCKVSQFIHL	YLMGCNRYFW	LCEGIYHLTL	240
IVVAVFAEKQ	HLMWYFFLW	GFPLIPACIH	AIARSLYYND	NCWISSDTHL	LYIIHGPICA	300
ALLVNLFFLL	NIVRVLITKL	KVTHQAESNL	YMKAVRATLI	LVPLLGIEFV	LIPWRPEGKI	360
ABEVYDYIMH	ILMHFQGLLV	STIFCFNNGE	VQAILRRNWN	QYKIQFGNSF	SNSEALRSAS	420
YTVSTISDGP	GYSHDCPSEH	LNGKSIHDIE	NVLLKPENLY	N		

Seq ID NO: 338 DNA sequence  
Nucleic Acid Accession #: NM\_001795  
Coding sequence: 25-2379

50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
GCACGATCTG	TTCCTCCTGG	GAAGATGCAG	AGGCTCATGA	TGCTCCTCGC	CACATCGGGC	60
GCCTGCCTGG	GCCTGCTGGC	AGTGGCAGCA	GTGGCAGCAG	CAGGTGCTAA	CCCTGCCCAA	120
CGGGACACCC	ACAGCCTGCT	GCCCACCCAC	CGGCGCCAAA	AGAGAGATTG	GATTTGGAAC	180
CAGATGCACA	TTGATGAAGA	GAAAAACACC	TCACTTCCCC	ATCATGTAGG	CAAGATCAAG	240
TCAAGCGTGA	GTCGCAAGAA	TGCCAAGTAC	CTGCTCAAAG	GAGAATATGT	GGGCAAGGTC	300
TTCCGGGTCG	ATGCAGAGAC	AGGAGACGTG	TTCCGCCATTG	AGAGGCTGGA	CCGGGAGAAAT	360
ATCTCAGAGT	ACCACCTCAC	TGCTGTCAAT	GTGGACAAGG	ACACTGGTGA	AAACCTGGAG	420
ACTCCTTCCA	GCTTCACCAT	CAAAGTTCAT	GACGTGAACG	ACAACTGGCC	TGTGTTACAG	480
CATCGGTTGT	TCAATGCGTC	CGTGCCTGAG	TCGTGCGCTG	TGGGGACCTC	AGTCATCTCT	540
GTGACAGCAG	TGGATGCAGA	CGACCCCACT	GTGGGAGACC	ACGCCTCTGT	CATGTACCAA	600
ATCCTGAAGG	GGAAAGAGTA	TTTTGCCATC	GATAATTCTG	GACGTATTAT	CACAATAACG	660
AAAAGCTTGG	ACCGAGAGAA	GCAGGCCAGG	TATGAGATCG	TGGTGAAGC	GCGAGATGCC	720
CAGGGCCTCC	GGGGGGACTC	GGGCACGGCC	ACCGTGCTGG	TCACTCTGCA	AGACATCAAT	780
GACAACTTCC	CCTTCTTCAC	CCAGACCAAG	TACACATTTG	TCGTGCCTGA	AGACACCCGT	840
GTGGGCACCT	CTGTGGGCTC	TCTGTTTGT	GAGGACCCAG	ATGAGCCCCA	GAACCGGATG	900
ACCAAGTACA	GCATCTTGCG	GGCGGACTAT	CAGGACGCTT	TCACCATTTG	GACAAACCCC	960
GCCCAACAAC	AGGCCATCAT	CAAGCCCATG	AAGCCTCTGG	ATTATGAATA	CATCCAGCAA	1020
TACAGCTTCA	TCGTGAGGGC	CACAGACCCC	ACCATCGACC	TCCGATACAT	GAGCCCTCCC	1080
GCGGGAAACA	GAGCCCAGGT	CATTATCAAC	ATCACAGATG	TGGACGAGCC	CCCCATTTTC	1140
CAGCAGCCTT	TCTACCACTT	CCAGCTGAAG	GAAAACCAGA	AGAAGCCTCT	GATTGGCACA	1200
GTGCTGGCCA	TGGACCCTGA	TGCGGCTAGG	CATAGCATTTG	GATACTCCAT	CCGCAGGACC	1260
AGTGACAAGG	GCCAGTTCTT	CCGAGTCAAC	AAAAAGGGGG	ACATTTACAA	TGAGAAAAGAA	1320
CTGGACAGAG	AAGTCTACCC	CTGGTATAAC	CTGACTGTGG	AGGCCAAAGA	ACTGGATTCC	1380
ACTGGAACCC	CCACAGGAAA	AGAATCCATT	GTGCAAGTCC	ACATTGAAGT	TTTGGATGAG	1440
AATGACAATG	CCCCGGAGTT	TGCCAAGCCC	TACCAGCCCA	AAGTGTGTGA	GAACGCTGTC	1500
CATGGCCAGC	TGGTCCCTGCA	GATCTCCGCA	ATAGACAAGG	ACATAACACC	ACGAAACGTG	1560
AAGTTCAAAT	TCACCTTGAA	TACTGAGAAC	AACTTTACCC	TCACGGATAA	TCACGATAAC	1620
ACGGCCAACA	TCACAGTCAA	GTATGGGCAG	TTTGACCGGG	AGCATACCAA	GGTCCACTTC	1680
CTACCCGTGG	TCATCTCAGA	CAATGGGATG	CCAATGTCGA	CGGGCACCAG	CACGCTGACC	1740
GTGGCCGTGT	GCAAGTGCAA	CGAGCAGGGC	GAGTTCACCT	TCTGCGAGGA	TATGGCCGCC	1800
CAGGTGGGCG	TGAGCATCCA	GGCAGTGGTA	GCCATCTTAC	TCTGCATCCT	CACCATCACA	1860
GTGATCACCC	TGCTCATCTT	CCTGCGGCGG	CGGCTCCGGA	AGCAGGCCCG	CGCGCACGGC	1920
AAGAGCGTGC	CGGAGATCCA	CGAGCAGCTG	GTCACTTACG	ACGAGGAGGG	CGGCGGCGAG	1980
ATGGACACCA	CCAGCTACGA	TGTGTCGGTG	CTCAACTCGG	TGCGCCGCGG	CGGGGCCAAG	2040
CCCCCGCGGC	CGCGCTGGA	CGCCCGGCTT	TCCCTCTATG	CGCAGGTGCA	GAAGCCACCG	2100
AGGCACGCGC	CTGGGGCACA	CGGAGGGCCC	GGGGAGATGG	CAGCCATGAT	CGAGGTGAAG	2160
AAGGACGAGG	CGGACACGA	CGGCGACGGC	CCCCCTTACG	ACACGCTGCA	CATCTACGGC	2220
TACGAGGGCT	CCGAGTCCAT	AGCCGAGTCC	CTCAGCTCCC	TGGGCACCGA	CTCATCCGAC	2280



TCTGACGTGG ATTACGACTT CCTTAACGAC TGGGGACCCA GGTTTAAGAT GCTGGCTGAG 2340  
 CTGTACGGCT CGGACCCCGG GGAGGAGCTG CTGTATTAGG CGGCCGAGGT CACTCTGGGC 2400  
 CTGGGGACCC AAACCCCTG CAGCCCAGGC CAGTCAGACT CCAGGCACCA CAGCCTCCAA 2460  
 5 AAATGGCAGT GACTCCCCAG CCCAGCACCC CTTCCTCGTG GGTCCCAGAG ACCTCATCAG 2520  
 CCTTGGGATA GCAAACTCCA GGTTCCTGAA ATATCCAGGA ATATATGTCA GTGATGACTA 2580  
 TTCTCAAATG CTGGCAAATC CAGGCTGGTG TTCTGTCTGG GCTCAGACAT CCACATAACC 2640  
 CTGTCACCCA CAGACCGCCG TCTAACTCAA AGACTTCCTC TGGCTCCCCA AGGCTGCAAA 2700  
 GCAAAACAGA CTGTGTTTAA CTGCTGCAGG GTCTTTTTCT AGGGTCCCTG AACGCCCTGG 2760  
 10 TAAGGCTGGT GAGGTCTCTG TGCTATCTG CCTGGAGGCA AAGGCCTGGA CAGCTTGACT 2820  
 TGTGGGGCAG GATTCTCTGC AGCCCATTC CAAAGGAGAC TGACCATCAT GCCCTCTCTC 2880  
 GGGAGCCCTA GCCCTGCTCC AACTCCATAC TCCACTCCAA GTGCCCCACC ACTCCCCAAC 2940  
 CCCTCTCCAG GCCTGTCAAG AGGGAGGAAG GGGCCCCATG GCAGCTCCTG ACCTTGGGTC 3000  
 CTGAAGTGAC CTCCTGGGCC TGCCATGCCA GTAACGTGTC TGTACTGAGC ACTGAACCAC 3060  
 ATTACAGGAA ATGCTTATTA AACCTTGAAG CAACTGTGAA TTCATTCTGG AGGGGCAGTG 3120  
 15 GAGATCAGGA GTGACAGATC ACAGGGTGAG GGCCACCTCC ACACCCACCC CCTCTGGAGA 3180  
 AGGCCCTGGA GAGCTGAGAC CTTGCTTTGA GACTCCTCAG CACCCCTCCA GTTTTGCCCTG 3240  
 AGAAGGGGCA GATGTTCCCG GAGATCAGAA GACGTCTCCC CTTCTCTGCC TCACCTGGTC 3300  
 GCCAATCCAT GCTCTCTTTC TTTTCTCTGT CTACTCCTTA TCCCTTGGTT TAGAGGAACC 3360  
 20 CAAGATGTGG CCTTAGCAAA AACTGACAAT GTCCAAACCC ACTCATGACT GCATGACGGA 3420  
 GCCGAGCATG TGTCTTTACA CCTCGCTGTT GTACATCTC AGGGAAGTGA CCCTCAGGCA 3480  
 CACCTTGCA AAGGAAGGCC CTGCCCTGTT CAACCTCTGT GGTCACCCAT GCATCATTC 3540  
 ACTGGAACGT TTTACTGCAA ACACACCTTG GAGAAGTGGC ATCAGTCAAC AGAGAGGGGC 3600  
 AGGGAAGGAG ACACCAAGCT CACCCTTCGT CATGGACCGA GGTTCCTACT CTGGCAAAGC 3660  
 25 CCCTCACACT GCAAGGGATT GTAGATAACA CTGACTTGT TGTTTTAAAC AATAACTAGC 3720  
 TTCTTATAAT GATTTTTTTA CTAATGATAC TTACAAGTT CTAGCTCTCA CAGACATATA 3780  
 GAATAAGGGT TTTTGCAATA TAAGCAGGTT GTTATTTAGG TTAACAATAT TAATTACAGT 3840  
 TTTTGTAGTG GAAAAACAAT TCCTGTAACC TTCTATTTTC TATAATTGTA GTAATTGCTC 3900  
 TACAGATAAT GTCTATATAT TGGCCAAACT GTGTCATGAC AAGTACTGTA TTTTTTTATA 3960  
 30 CCTAAATAAA GAAAAATCTT TAGCCTGGGC AACAAAAAAA

Seq ID NO: 339 Protein sequence  
 Protein Accession #: NP\_001786

1 11 21 31 41 51  
 | | | | |  
 35 MQRLMMLLAT SGACLGLLAV AAVAAAGANP AQRDTHSLLP THRRQKRDWI WNQMHHIDEK 60  
 NTSLPHHVKG IKSSVSRKNA KYLLKGEYVG KVFRVDAETG DVFAIERLDR ENISEYHLTA 120  
 VIVDKDTGEN LETPSSFTIK VHDVNDNWPV FTHRLFNASV PESSAVGTSV ISVTAVDADD 180  
 40 PTVGDHASVM YQILKGKEYF AIDNSGRIIT ITKSLDREKQ ARYEIVVEAR DAQGLRGDSG 240  
 TATVLVTLQD INDNFFPFTQ TKYTFVVPED TRVGTSGVSL FVEDPDEPQN RMTKYSILRG 300  
 DYQDAFTIET NPAHNEGIIK PMKPLDYEYI QQYSFIVEAT DPTIDLRYMS PPAGNRAQVI 360  
 INITDVDEPP IFQQPFYHFQ LKENQKKPLI GTVLAMPDPA ARHSIGYSIR RTSKQGFRR 420  
 VTKKGDYNE KELDREVYPW YNLTVAEKEL DSTGTPTGKE SIVQVHIEVL DENDNAPEFA 480  
 45 KPYQPKVCEN AVHQLVLQI SAIDKDIPTP NVKFKFTLNT ENNFTLTDNH DNTANITVKY 540  
 GQFDREHTKV HFLPVVISDN GMPSTRGTST LTVAVCKCNE QGEFTFCEDM AAQVGVSIQA 600  
 VVAILLCILT ITVITLLIFL RRLRLKQARA HGKSVPEIHE QLVITYDEEGG GEMDITSYDV 660  
 SVLNSVRRGG AKPPRPALDA RPSLYAQVQK PPRHAPGAHG GPGEAMAMIE VKKDEADHDG 720  
 DGPPYDTLHI YGYEGSESIA ESLSSLGTDS SDSVDYDFL NDWGPFRFKML AELYGSDPRE 780  
 50 ELLY

Seq ID NO: 340 DNA sequence  
 Nucleic Acid Accession #: NM\_003088  
 Coding sequence: 112-1593

1 11 21 31 41 51  
 | | | | |  
 55 GCGGAGGGTG CGTGCGGGCC GCGGCAGCCG AACAAAGGAG CAGGGGCGCC GCCGCAGGGA 60  
 CCCGCCACCC ACCTCCCGGG GCGGCGCAGC GGCCTCTCGT CTACTGCCAC CATGACCGCC 120  
 60 AACGGCACAG CCGAGGCGGT GCAGATCCAG TTCGGCCTCA TCAACTGCGG CAACAAGTAC 180  
 CTGACGGCCG AGGCGTTCGG GTTCAAGGTG AACGCGTCCG CCAGCAGCCT GAAGAAGAAG 240  
 CAGATCTGGA CGCTGGAGCA GCCCCCTGAC GAGGCGGGCA GCGCGGCCGT GTGCCTGCGC 300  
 AGCCACCTGG GCCGCTACCT GCGGCGCGAC AAGGACGGCA ACGTGACCTG CGAGCGCGAG 360  
 GTGCCCCGTC CCGACTGCCG TTTCTCATC GTGGCGCACG ACGACGGTCG CTGGTCTGCTG 420  
 65 CAGTCCGAGG CGCACCGGCG CTAAGTCCGG GGCACGAGG ACCGCTGTC CTGCTTCGCG 480  
 CAGACGGTGT CCCCCCGCGA GAGCTGGAGC GTGCACATCG CCATGCACCC TCAGGTCAAC 540  
 ATCTACAGTG TCACCCGTAA GCGCTACGCG CACCTGAGCG CGCGGCCGGC CGACGAGATC 600  
 GCCGTGGACC GCGACGTGCC CTGGGCGGTC GACTCGCTCA TCACCTCGC CTTCCAGGAC 660  
 70 CAGCGCTACA GCGTGCAGAC CGCCGACCAC CGCTTCTGTC GCCACGACGG GCGCCTGGTG 720  
 GCGCGCCCCG AGCCGGCCAC TGGCTACACG CTGGAGTTCC GCTCCGGCAA GGTGGCCTTC 780  
 CGCGACTGCG AGGGCCGTTA CCTGGCGCCG TCGGGGCCCA GCGGCACGCT CAAGGCGGGC 840  
 AAGGCCACCA AGGTGGGCAA GGACGAGCT TTTGCTCTGG AGCAGAGCTG CGCCAGGTC 900  
 GTGCTGCAGG CGGCCAACGA GAGGAACGTG TCCACGCGCC AGGGTATGGA CCTGTCTGCC 960  
 75 AATCAGGACG AGGAGACCGA CCAGGAGACC TTCCAGCTGG AGATCGACCG CGACACCAA 1020  
 AAGTGTGCC TCCGTACCCA CACGGGCAAG TACTGGACGC TGACGGCCAC CGGGGGCGTG 1080  
 CAGTCCACCG CCTCCAGCAA GAATGCCAGC TGCTACTTTG ACATCGAGTG GCGTGACCGG 1140  
 CGCATCACAC TGAGGGCGTC CAATGGCAAG TTTGTACCT CCAAGAAGAA TGGGCAGCTG 1200  
 GCGGCTCGG TGGAGACAGC AGGGGACTCA GAGCTCTTCC TCATGAAGCT CATCAACCGC 1260  
 80 CCCATCATCG TGTTCGCGCG GGAGCATGGC TTCATCGGCT GCCGCAAGGT CACGGGCACC 1320  
 CTGGACGCCA ACCGCTCCAG CTATGACGTC TTCCAGCTGG AGTTCAACGA TGGCGCTAC 1380  
 AACATCAAAG ACTCCACAGG CAAATACTGG ACGGTGGGCA GTGACTCCGC GGTCAACAGC 1440  
 AGCGGCGACA CTCCTGTGGA CTCTTCTTTC GAGTTCTGCG ACTATAACAA GGTGGCCATC 1500  
 AAGGTGGGCG GCGCTACCT GAAGGGCGCG CACGCAGGCG TCCTGAAGGC CTGGCGGAA 1560  
 85 ACCGTGGACC CCGCTCGCT CTGGGAGTAC TAGGGCCGGC CCGTCTTCC CCGCCCCCTG 1620  
 CCACATGGCG GCTCTGCCA ACCCTCCCTG CTAACCCCTT CTCCGCCAGG TGGGCTCCAG 1680  
 GCGGGGAGGC AAGCCCCCTT GCCTTTCAA CTGGAAACCC CAGAGAAAAC GGTGCCCCCA 1740  
 CCTGTGCCCC CTATGACTC CCCACTCTCC CCTCCGCCG GGTTCCTTAC TCCCCTCGGG 1800



5  
10  
15

```

TCAGCGGCTG CGGCCTGGCC CTGGGAGGGA TTTTCAGATGC CCCTGCCCTC TTGTCTGCCA 1860
CGGGGCGAGT CTGGCACCTC TTTCTTCTGA CCTCAGACGG CTCTGAGCCT TATTTCTCTG 1920
GAAGCGGCTA AGGGACGGTT GGGGCTGGG AGCCCTGGGC GTGTAGTGTA ACTGGAATCT 1980
TTTGCCTCTC CCAGCCACCT CCTCCCAGCC CCCCAGGAGA GCTGGGCACA TGTCCCAAGC 2040
CTGTCACTGG CCCTCCCTGG TGCAGTGTCC CCGAAACCCC TGCTTGGGAA GGAAGCTGT 2100
CGGGAGGGCT AGGACTGACC CTGTGGTGT TTTTTTGGGT GGTGGCTGGA AACAGCCCCT 2160
CTCCACAGTG GGAGAGGCTC AGCCTGGCTC CTTCCCTGG AGCGGCAGGG CGTGACGGCC 2220
ACAGGGTCTG CCCGCTGCAC GTTCTGCCAA GGTGGTGGTG GCGGGCGGGT AGGGGTGTGG 2280
GGGCGGTCTT CCTCCTGTCT CTTTCTTTT ACCCTAGCCT GACTGGAAGC AGAAAATGAC 2340
CAAATCAGTA TTTTTTTTAA TGAAATATTA TTGCTGGAGG CGTCCCAGGC AAGCCTGGCT 2400
GTAGTAGCGA GTGATCTGGC GGGGGGCGTC TCAGCACCTC CCCCAGGGGG TGCATCTCAG 2460
CCCCCTCTTT CCGTCCCTCC CGTCCAGCCC CAGCCCTGGG CTTGGGCTGC CGACACCTGG 2520
GCCAGAGCCC CTGCTGTGAT TGGTGTCTCC TGGGCTCTCC GGTGGATGA AGCCAGGCGT 2580
CGCCCCCTCC GGGAGCCCTG GGGTGAGCCG CCGGGGCCCC CTGCTGCCA GCCTCCCCCG 2640
TCCCCAACAT GCATCTCACT CTGGGTGTCT TGGTCTTTTA TTTTGTGTA GTGTCAATTG 2700
TATAACTCTA AACGCCCATG ATAGTAGCTT CAAACTGGAA ATAGCGAAAT AAAATAACTC 2760
AGTCTGC

```

20 Seq ID NO: 341 Protein sequence  
Protein Accession #: NP\_003079

25  
30

```

1 11 21 31 41 51
| | | | |
MTANGTAEAV QIQFGLINCG NKYLTAEAFG FKVNASASSL KKKQIWTLEQ PPDEAGSAAV 60
CLRSHLGRYL AADKDGNTVC EREVPGPCR FLIVAHDDGR WSLQSEAHRR YFGGTEDRLS 120
CFAQTVSPA E KWSVHIAMHP QVNIYSVTRK RYAHLSARPA DEIAVDRDVP WGVDSLITLA 180
FQDQRYSVQT ADHRFLRHDG RLVARPEPAT GYTLEFRSGK VAFRDCEGRY LAPSGPSGTL 240
KAGKATKVGK DELFALEQSC AQVVLQAANE RNVSTRQGM L SANQDEETD QETFOLEIDR 300
DTKKCAFRTH TGKYWTLTAT GGVQSTASSK NASCYFDIEW RDRRITLRAS NGKFVTSKKN 360
GQLAASVETA GDSEFLMKL INRPIIVFRG EHGFIGCRKV TGTLDANRSS YDVFQLEFND 420
GAYNIKDSTG KYWTVGSDSA VTSSGDTFVD FFFEFCDYNK VAIKVGGRYL KGDHAGVLKA 480
SAETVDPASL WEY

```

35 Seq ID NO: 342 DNA sequence  
Nucleic Acid Accession #: FGENESH predicted  
Coding sequence: 660..1705

40  
45  
50  
55  
60  
65  
70

```

1 11 21 31 41 51
| | | | |
CGCTCCGCAC ACATTTCTCTG TCGCGGCCTA AGGGAAACTG TTGGCCGCTG GGCCCGCGGG 60
GGGATTCTTG GCAGTTGGGG GGTCCGTCGG GAGCGAGGGC GGAGGGGAAG GGAGGGGGAA 120
CCGGGTTGGG GAAGCCAGCT GTAGAGGGCG GTGACCGCGC TCCAGACACA GCTCTGCGTC 180
CTCGAGCGGG ACAGATCCAA GTTGGGAGCA GCTCTGCGTG CCGGGCCTCA GAGAATGAGG 240
CCGGCGTTTG CCCTGTGCCT CCTCTGGCAG GCGCTCTGGC CCGGGCCGGG CGGCGGCGAA 300
CACCCCACTG CCGACCGTGC TGGTGTCTCG GCCTCGGGGG CCTGCTACAG CCTGCACCAC 360
GCTACCATGA AGCGGCAGGC GGCCGAGGAG GCCTGCATCC TCGAGGTGG GCGCTCAGC 420
ACCGTGCGTG CCGGCGCCGA GCTGCGCGCT GTGCTCGCGC TCCTGCGGGC AGGCCAGGG 480
CCCGGAGGGG GCTCCAAAGA CCTGCTGTTC TGGGTGCGAC TGGAGCGCAG GCGTTCACAC 540
TGCAACCTGG AGAACGAGCC TTTGCGGGGT TTCTCCTGGC TGTCTCCGA CCGCGCGGGT 600
CTCGAAAGCG ACACGCTGCA GTGGGTGGAG GAGCCCCAAC GCTCCTGCAC CGCGCGGAGA 660
TGCGCGGTAC TCCAGGCCAC CGGTGGGTCT GAGCCCGCAG CTGGAAGGAG ATGCGATGCC 720
ACCTGCGCGC CAACGGCTAC CTGTGCAAGT ACCAGTTTGA GGTCTTGTGT CCTGCGCCGC 780
GCCCCGGGGC CGCCTCTAAC TTGAGCTATC GCGCGCCCTT CCAGCTGCAC AGCGCCGCTC 840
TGGACTTCAG TCCACCTGGG ACCGAGGTGA GTGCGCTCTG CCGGGGACAG CTCCCGATCT 900
CAGTTACTTG CATCGCGGAC GAAATCGGCG CTCGCTGGGA CAAACTCTCG GGCGATGTGT 960
TGTGTCCCTG CCGCGGGAGG TACCTCCGTG CTGGCAAATG CGCAGAGCTC CCTAACTGCC 1020
TAGACGACTT GGGAGGCTTT GCCTGCGAAT GTGCTACGGG CTTTCGAGCTG GGAAGGACG 1080
CGCGCTCTTG TGTGACAGT GGGGAAGGAC AGCCGACCTT TGGGGGGACC GGGGTGCCCA 1140
CCAGGCGCCC GCGGCGCACT GCAACCAGCC CCGTGCCGCA GAGAACATGG CCAATCAGGG 1200
TCGACGAGAA GCTGGGAGAG ACACCACTTG TCCCTGAACA AGACAATTCA GTAACATCTA 1260
TTCTGAGAT TCCTCGATGG GGATCACAGA GCACGATGTC TACCCTTCAA ATGTCCCTTC 1320
AAGCCGAGTC AAAGGCCACT ATCACCCTAT CAGGGAGCGT GATTTCCAAG TTTAATTCTA 1380
CGACTTCCTC TGCCACTCCT CAGGCTTTTG ACTCCTCCTC TGCCGTGGTC TTCATATTG 1440
TGAGCACAGC AGTAGTAGT TTGGTGATCT TGACCATGAC AGTACTGGGG CTTGTCAAGC 1500
TCTGCTTTCA CGAAAGCCCC TCTTCCAGC CAAGGAAGGA GTCTATGGGC CCGCCGGGCC 1560
TGGAGAGTGA TCCTGAGCCC GCTGCTTTGG GCTCCAGTTC TGCACATTGC ACAAACAATG 1620
GGGTGAAAGT CCGGGACTGT GATCTGCGGG ACAGAGCAGA GGGTGCCTTG CTGGCGGAGT 1680
CCCCTCTTGG CTCTAGTGAT GCATAG

```

75 Seq ID NO: 343 Protein sequence  
Protein Accession #: FGENESH predicted

75  
80  
85

```

1 11 21 31 41 51
| | | | |
MGKDFMTKTP KAFATKAKID KWDLIKLSKF CTAKETIIRV NSQPTDWQKT FAIYPSDKGV 60
IARIYKELEQ IYKKKKPTKT LRTHFLSRPK GNCWPLGPRG DSWQLGGPSG ARAEGKGGT 120
GLGKPAVEGG DRAPDTALRP RAGQIQVSS SACGASENEA GVRVPVPLAG ALARAGRRRT 180
PHCRPCWLLG LGGLLQAPAPR YHEAAGGRGG LHPARWGAQH RACGRRAARC ARAPAGRPR 240
RRGLQRPVAVL GRTGAQAFPL HPGERAFAGF LLAVLRPRRS RKRHAAVGGG APTLLHRAEM 300
RGTPGHRWGR ARSWKEMRCH LRANGYLCKY QFEVLCAPAPR PGAASNLSYR APFQLHSAAL 360
DFSPPGTEVS ALCRGQLPIS VTCIADEIGA RWDKLSGDVL CPCPGRYLRA GKCAELPNCL 420
DDLGGFACEC ATGFELGKDG RSCVTSSEGG TMLGGTGVPT RRPATATSP VPQRTWPIRV 480
DEKLGETPLV PEIQDNSVTSI PEIPRWGSQS TMSTLQMSLQ AESKATITPS GSVISKFNST 540
TSSATPQAFD SSSAVVFIFV STAVVVLVIL TMTVLGLVKL CFHESPSSQP RKESMGPPGL 600
ESDPEPAALG SSSAHCTNNG VKVGDCDLRD RAEGALLAES PLGSSDA

```



Seq ID NO: 344 DNA sequence  
Nucleic Acid Accession #: NM\_012072  
Coding sequence: 149-2107

5  
0  
5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
AAAGCCCTCA	GCCTTTGTGT	CCTTCTCTGC	GGCGGAGTGG	CTGCAGCTCA	CCCCTCAGCT	60
CCCCTTGGGG	CCCAGCTGGG	AGCCGAGATA	GAAGCTCCTG	TCGCCGCTGG	GCTTCTCGCC	120
TCCCGCAGAG	GGCCACACAG	AGACCGGGAT	GGCCACCTCC	ATGGGCCTGC	TGCTGCTGCT	180
GCTGCTGCTC	CTGACCCAGC	CCGGGGCGGG	GACGGGAGCT	GACACGGAGG	CGGTGGTCTG	240
CGTGGGGACC	GCCTGTCTAC	CGGCCCACTC	GGGCAAGCTG	AGCGCTGCCG	AGGCCCAGAA	300
CCACTGCAAC	CAGAACGGGG	GCAACCTGGC	CACTGTGAAG	AGCAAGGAGG	AGGCCCAGCA	360
CGTCCAGCGA	GTACTGGCCC	AGCTCCTGAG	GCGGGAGGCA	GCCCTGACGG	CGAGGATGAG	420
CAAGTTCTGG	ATTGGGCTCC	AGCGAGAGAA	GGGCAAGTGC	CTGGACCCTA	GTCTGCCGCT	480
GAAGGGCTTC	AGCTGGGTGG	GCGGGGGGGA	GGACACGCCT	TACTCTAACT	GGCACAAGGA	540
GCTCCGGAAC	TCGTGCATCT	CCAAGCGCTG	TGTGTCTCTG	CTGCTGGACC	TGTCCAGGCC	600
GCTCCTTCCC	AACCGCCTGC	CCAAGTGGTC	TGAGGGCCCC	TGTGGGAGCC	CAGGCTCCCC	660
CGGAAGTAAC	ATTGAGGGCT	TCGTGTGCAA	GTTCAGCTTC	AAAGGCATGT	GCCGGCCTCT	720
GGCCCTGGGG	GGCCAGGTTC	AGGTGACCTA	CACCACCCCC	TTCCAGACCA	CCAGTTCTCT	780
CTTGAGGGCT	GTGCCCTTTG	CCTCTGCGGC	CAATGTAGCC	TGTGGGGAAG	GTGACAAGGA	840
CGAGACTCAG	AGTCATTATT	TCCTGTGCAA	GGAGAAGGCC	CCCGATGTGT	TCGACTGGGG	900
CAGCTCGGGC	CCCCTCTGTG	TCAGCCCCAA	GTATGGCTGC	AACTTCAACA	ATGGGGGCTG	960
CCACCAGGAC	TGCTTTGAAG	GGGGGGATGG	CTCCTTCCTC	TGCGGCTGCC	GACCAGGATT	1020
CCGGCTGCTG	GATGACCTGG	TGACCTGTGC	CTCTCGAAAC	CCTTGCAGCT	CCAGCCCATG	1080
TCGTGGGGGG	GCCACGTGCG	TCCTGGGACC	CCATGGGAAA	AACTACACGT	GCCGCTGCCC	1140
CCAAGGGTAC	CAGCTGGACT	CGAGTCAGCT	GGACTGTGTG	GACGTGGATG	AATGCCAGGA	1200
CTCCCCCTGT	GCCCAGGAGT	GTGTCAACAC	CCCTGGGGGC	TTCCGCTGCG	AATGCTGGGT	1260
TGGCTATGAG	CCGGGCGGTC	CTGGAGAGGG	GGCCTGTCAG	GATGTGGATG	AGTGTGCTCT	1320
GGGTGCTCG	CCTTGCGCCC	AGGGCTGCAC	CAACACAGAT	GGCTCATTTT	ACTGCTCCTG	1380
TGAGGAGGGC	TACGTCTCTG	CCGGGGAGGA	CGGGACTCAG	TGCCAGGACG	TGGATGAGTG	1440
TGTGGGCCCC	GGGGGCCCCC	TCTGCGACAG	CTTGTGCTTC	AACACACAAG	GGTCCTTCCA	1500
CTGTGGCTGC	CTGCCAGGCT	GGGTGCTGGC	CCCAAATGGG	GTCTCTTGCA	CCATGGGGCC	1560
TGTGTCTCTG	GGACCACCAT	CTGGGCCCCC	CGATGAGGAG	GACAAAGGAG	AGAAAGAAGG	1620
GAGCACCGTG	CCCCGCGCTG	CAACAGCCAG	TCCCACAAGG	GGCCCCGAGG	GCACCCCCAA	1680
GGCTACACCC	ACCACAAGTA	GACCTTCGCT	GTCATCTGAC	GCCCCCATCA	CATCTGCCCC	1740
ACTCAAGATG	CTGGCCCCCA	GTGGGTCTCT	AGGCGTCTGG	AGGGAGCCCA	GCATCCATCA	1800
CGCCACAGCT	GCCTCTGGCC	CCCAGGAGCC	TGCAGGTGGG	GACTCCTCCG	TGGCCACACA	1860
AAACAACGAT	GGCACTGACG	GGCAAAAGCT	GCTTTTATTC	TACATCCTAG	GCACCGTGGT	1920
GGCCATCCTA	CTCCTGCTGG	CCCTGGCTCT	GGGGCTACTG	GTCTATCGCA	AGCGGAGAGC	1980
GAAGAGGGAG	GAGAAGAAGG	AGAAGAAGCC	CCAGAATGCG	GCAGACAGTT	ACTCCTGGGT	2040
TCCAGAGCGA	GCTGAGAGCA	GGGCCATGGA	GAACCAGTAC	AGTCCGACAC	CTGGGACAGA	2100
CTGCTGAAAG	TGAGGTGGCC	CTAGAGACAC	TAGAGTCACC	AGCCACCATC	CTCAGAGCTT	2160
TGAACTCCCC	ATTCCAAAGG	GGCACCCACA	TTTTTTTGAA	AGACTGGACT	GGAACTCTAG	2220
CAAACAATTG	TAAGTCTCCT	CCTTAAAGGC	CCCTTGGAAC	ATGCAGGTAT	TTTCTACGGG	2280
TGTTTGATGT	TCCTGAAGTG	GAAGCTGTGT	GTTGGCGTGC	CACGGTGGGG	ATTTCTGTAC	2340
TCTATAATGA	TTGTTACTCC	CCCTCCCTTT	TCAAATTCCA	ATGTGACCAA	TTCCGGATCA	2400
GGGTGTGAGG	AGGCTGGGGC	TAAGGGGCTC	CCCTGAATAT	CTTCTCTGCT	CACTTCCACC	2460
ATCTAAGAGG	AAAAGGTGAG	TTGCTCATGC	TGATTAGGAT	TGAAATGATT	TGTTTCTCTT	2520
CCTAGGATGA	AAACTAAATC	AATTAATTAT	TCAATTAGGT	AAGAAGATCT	GGTTTTTTGG	2580
TCAAAGGGAA	CATGTTCCGA	CTGGAAACAT	TTCTTTACAT	TTGCATTCCT	CCATTTCCGC	2640
AGCACAAAGT	TTGCTAAATG	TGATACTGTT	GACATCCTCC	AGAATGGCCA	GAAGTGCAAT	2700
TAACCTCTTA	GGTGGCAAAG	AGGCAGGAAG	TGCCTCTTTA	GTTCTTACAT	TTCTAATAGC	2760
CTTGGGTTTA	TTTGCAAAGG	AAGCTTGAAA	AATATGAGAA	AAGTTGCTTG	AAGTGCATTA	2820
CAGGTGTTTG	TGAAGTCACA	TAATCTACGG	GGCTAGGGCG	AGAGAGGCCA	GGGATTTGTT	2880
CACAGATACT	TGAATTAATT	CATCCAAATG	TACTGAGGTT	ACCACACACT	TGACTACGGA	2940
TGTGATCAAC	ACTAACAAGG	AAACAATTC	AAGGACAACC	TGTCTTTGAG	CCAGGGCAGG	3000
CCTCAGACAC	CCTGCCTGTG	GGCCCGCCTC	CACTTCATCC	TGCCCGGAAT	GCCAGTGCTC	3060
CGAGCTCAGA	CAGAGGAAGC	CCTGCAGAAA	GTTCCATCAG	GCTGTTTCCT	AAAGGATGTG	3120
TGAACGGGAG	ATGATGCACT	GTGTTTGAAG	AGTTGTCATT	TAAAGCATT	TTAGCACAGT	3180
TCATAGTCCA	CAGTTGATGC	AGCATCCTGA	GATTTTAAAT	CCTGAAGTGT	GGGTGGCGCA	3240
CACACCAAGT	AGGGAGCTAG	TCAGGCAGTT	TGCTTAAGGA	ACTTTTGTTC	TCTGTCTCTT	3300
TTCTTTAAAA	TTGGGGGTAA	GGAGGGAAAG	AAGAGGGAAA	GAGATGACTA	ACTAAATCA	3360
TTTTTACAGC	AAAACTGCT	CAAAGCCATT	TAAATTATAT	CCTCATTTTA	AAAGTTACAT	3420
TTGCAAAATAT	TTCTCCCTAT	GATAATGCGA	TCGATAGTGT	GCACTCTTTC	TCTCTCTCTC	3480
TCTCTCTCAC	ACACACACAC	ACACACACAC	ACACACACAC	AGAGACACGG	CACCATCTCTG	3540
CCTGGGGCAC	TGGAACACAT	TCCTGGGGGT	CACCGATGGT	CAGAGTCACT	AGAAGTTACC	3600
TGAGTATCTC	TGGGAGGCCT	CATGTCTCCT	GTGGGCTTTT	TACCACCACT	GTGCAGGAGA	3660
ACAGACAGAG	GAAATGTGTC	TCCCTCCAAG	GGCCCAAAGC	CTCAGAGAAA	GGGTGTTTCT	3720
GGTTTTGCCT	TAGCAATGCA	TCGGTCTCTG	AGGTGACACT	CTGGAGTGGT	TGAAGGGCCA	3780
CAAGGTGCAG	GGTTAATACT	CTTGCCAGTT	TTGAAATATA	GATGCTATGG	TTTCAATTTG	3840
TTTTAATAGA	AAACTAAAGG	GGCAGGGGAA	GTGAAAGGAA	AGATGGAGGT	TTTGTGCGGC	3900
TCGATGGGGC	ATTTGGAAGT	TCTTTTAAAA	GTCTCTCTCAT	GGTCTCCAGT	TTTCAATTTG	3960
AACTCTGGTG	TTTAACACTT	AAGGGAGACA	AAGGCTGTGT	CCATTTGGCA	AAACTTCCTT	4020
GGCCACGAGA	CTCTAGGTGA	TGTGTGAAGC	TGGGCAGTCT	GTGGTGTGGA	GAGCAGCCAT	4080
CTGTCTGGCC	ATTCAGAGGA	TTCTAAAGAC	ATGGCTGGAT	GCGCTGCTGA	CCAACATCAG	4140
CACTTAAATA	AATGCAAATG	CAACATTTCT	CCCTCTGGGC	CTTGAAAATC	CTTGCCCTTA	4200
TCATTTGGGG	TGAAGGAGAC	ATTTCTGTCC	TTGGCTTCCC	ACAGCCCCAA	CGCAGTCTGT	4260
GTATGATTCC	TGGGATCCAA	CGAGCCCTCC	TATTTTTCACA	GTGTTCTGAT	TGCTCTCACA	4320
GCCCAGGCCC	ATCGTCTGTT	CTCTGAATGC	AGCCCTGTTC	TCAACAACAG	GGAGGTCATG	4380
GAACCCCTCT	GTGGAACCCA	CAAGGGGAGA	AATGGGTGAT	AAAGAATCCA	GTTCCTCAAA	4440
ACCTTCCCTG	GCAAGGCTGG	TCCCTCTCCT	GCTGGGTGGT	GCTTTCTCTT	GCACACCACT	4500
CCCACCACGG	GGGGAGAGCC	AGCAACCCAA	CCAGACAGCT	CAGGTTGTGC	ATCTGATGGA	4560
AACCACTGGG	CTCAACACAG	TGCTTTATTC	TCCTGTTTAT	TTTTGCTGTT	ACTTTGAAGC	4620
ATGGAAATTC	TTGTTGGGG	GATCTTGGGG	CTACAGTAGT	GGGTAAACAA	ATGCCACCG	4680
GCCAAGAGGC	CATTAACAAA	TCGTCTTTGT	CCTGAGGGGC	CCCAGCTTGC	TGGGGCGTGG	4740
CACAGTGGGG	AATCCAAGGG	TCACAGTATG	GGGAGAGGTG	CACCTGCCA	CCTGCTAACT	4800



5  
10  
15  
20  
25  
30

```

TCTCGCTAGA CACAGTGTTC CTGCCCAGGT GACCTGTTCA GCAGCAGAAC AAGCCAGGGC 4860
CATGGGGACG GGGGAAGTTT TCACCTGGAG ATGGACACCA AGACAATGAA GATTGTGTGT 4920
CCAAATAGGT CAATAATTCT GGGAGACTCT TGGAAAAAAC TGAATATATT CAGGACCAAC 4980
TCTCTCCCTC CCCTCATCCC ACATCTCAAA GCAGACAATG TAAAGAGAGA ACATCTCACA 5040
CAGCCAGCTC GCCATGCCTA CTCATTCTGT AATTTTCAGGT GCCATCACTG CTCTTTCTTT 5100
CTTCTTTGTC ATTTGAGAAA GGATGCAGGA GGACAATTCC CACAGATAAT CTGAGGAATG 5160
CAGAAAAACC AGGGCAGGAC AGTTATCGAC AATGCATTAG AACTTGGTGA GCATCCTCTG 5220
TAGAGGGACT CCACCCCTGC TCAACAGCTT GGCTTCCAGG CAAGACCAAC CACATCTGGT 5280
CTCTGCCTTC GGTGGCCAC ACACCTAAGC GTCATCGTCA TTGCCATAGC ATCATGATGC 5340
AACACATCTA CGTGTAGCAC TACGACGTTA TGTTTGGGTA ATGTGGGGAT GAACTGCATG 5400
AGGCTCTGAT TAAGGATGTG GGAAGTGGG CTGCGGTCAC TGTCGGCCTT GCAAGGCCAC 5460
CTGGAGGCCT GTCTGTTAGC CAGTGGTGGG GGAGCAAGGC TTCAGGAAGG GCCAGCCACA 5520
TGCCATCTTC CCTGCGATCA GGCAAAAAAG TGGAATTAAA AAGTCAAACC TTTATATGCA 5580
TGTGTTATGT CCATTTTGCA GGATGAACAT AGTTTAAAG AATTTTTTTT TCTCTTCAAG 5640
TTGCTTTGTC TTTTCCATCC TCATCACAAAG CCCTGTGTTG AGTGTCTTAT CCCTGAGCAA 5700
TCTTTCGATG GATGGAGATG ATCATTAGGT ACTTTTGTTC CAACCTTTAT TCCTGTAAAT 5760
ATTTCTGTGA AAACAGGAG AACAGAGATG AGATTGACA AAAAAAATT GAATTAAAAA 5820
TAACACAGTC TTTTAAAC TAACATAGGA AAGCCTTTCC TATTATTTCT CTCTTAGCT 5880
TCTCCATTGT CTAAATCAGG AAAACAGGAA AACACAGCTT TCTAGCAGCT GCAAAATGGT 5940
TTAATGCCCC CTACATATTT CCATCACCTT GAACAATAGC TTTAGCTTGG GAATCTGAGA 6000
TATGATCCCA GAAACATCT GTCTCTACTT CGGCTGCAAA ACCCATGGTT TAAATCTATA 6060
TGGTTTGTGC ATTTTCTCAA TAAAAAAT AGATGATAAT CCGAATTCTC CATATATTCA 6120
CTAATCAAAG AACTATTTT CATACTAGAT TCCTGAGACA AATACTCACT GAAGGGCTTG 6180
TTTAAAAATA AATTGTGTTT TGGTCTGTC TTGTAGATAA TGCCCTTCTA TTTAGGTAG 6240
AAGCTCTGGA ATCCCTTTAT TGTGCTGTG CTCTTATCTG CAAGGTGGCA AGCAGTTCTT 6300
TTCAGCAGAT TTTGCCACT ATTCCTCTGA GCTGAAGTTC TTTGCATAGA TTTGGCTTAA 6360
GCTTGAATTA GATCCCTGCA AAGGCTTGCT CTGTGATGTC AGATGTAATT GTAAATGTCA 6420
GTAATCACTT CATGAATGCT AAATGAGAAT TTAATGTGT GTATTTCAAA 6480
TTTGTGAC TAATTCTGGA ATTACAAGAT TTCTATGCAG GATTTACCTT CATCCTGTGC 6540
ATGTTTCCCA AACTGTGAGG AGGGAAGGCT CAGAGATCGA GCTTCTCCTC TGAGTTCTAA 6600
CAAAATGGTG CTTTGAGGGT CAGCCTTTAG GAAGGTGCAG CTTTGTTGTC CTTTGAGCTT 6660
TCTGTTATGT GCCTATCCTA ATAACTCTT AAACACATT

```

Seq ID NO: 345 Protein sequence  
Protein Accession #: NP\_036204

40  
45  
50

```

1 11 21 31 41 51
| | | | |
MATSMGLLLL LLLLLTQPGA GTGADTEAVV CVGTACYTAH SGKLSAAEAQ NHCNQNGGNL 60
ATVKSKEEAQ HVQRVLAQLL RREAALTARM SKFWIGLQRE KGKCLDPSLP LKGFVWVGGG 120
EDTPYSNWHK ELRNSCISKR CVSLLDLSQ PLEPNRLPKW SEGPCGSPGS PGSNIEGFVC 180
KFSFKMCRP LALGGPGQVT YTPFQTSS SLEAVPFASA ANVACGEGDK DETQSHYFLC 240
KEKAPDVFDW GSSGPLCVSP KYGCNFFNNG CHQDCFEGGD GSFLCGCRPG FRLLDDLVTIC 300
ASRNPCSSSP CRGGATCVLG PHGKNYTCRC PQGYQLDSSQ LDCVDVDECQ DSPCAQEVCN 360
TPGGFRCECW VGYEPGGPGE GACQDVDECA LGRSPCAQGC TNTDGSFHCS CEEGYVLAGE 420
DGTQCQDVDE CVGPGGPLCD SLCFNTQDEA HCGCLPGWVL APNGVSCTMG PVS LGPPSGP 480
PDEEDKGEKE GSTVPRATA SPTRGPEGTP KATPTTSRPS LSSDAPITSA PLKMLAPSGS 540
SGVWREPSIH HATAASGPQE PAGGDSSVAT QNNDGTDGQK LLLFYILGTV VAILLLLLALA 600
LGLLVYRKRR AKREEKKEKK PQNAADSYSW VPERAESRAM ENQYSPTPGT DC

```

Seq ID NO: 346 DNA sequence  
Nucleic Acid Accession #: Z31560  
Coding sequence: <1-966

55  
60  
65  
70  
75

```

1 11 21 31 41 51
| | | | |
CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCCGGG CCCGCAGCAA 60
ACTTCGGGGG GCGGCGGCGG CAATCCACC GCGGCGGCGG CCGGCGGCAA CCAGAAAAAC 120
AGCCCGGACC GCGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCCG CGGGCAGCGG 180
CGCAAGATGG CCCAGGAGAA CCCCAGATG CACAACCTCG AGATCAGCAA GCGCCTGGGC 240
GCCGAGTGGG AACTTTTGTG GGAGACGGAG AAGCGGCCGT TCATCGACGA GGCTAAGCGG 300
CTGCGAGCGC TGCACATGAA GGAGCACCCG GATTATAAAT ACCGGCCCCG GCGGAAAAAC 360
AAGACGCTCA TGAAGAAGGA TAAGTACACG CTGCCCGGCG GGCTGCTGGC CCCCGGCGGC 420
AATAGCATGG CGAGCGGGGT CGGGGTGGG GCGGCGCTGG GCGGCGGCGT GAACCAGCGC 480
ATGGACAGTT ACGGCGACAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540
CAGCTGGGCT ACCCGCAGCA CCCGGGCCTC AATGCGCACG GCGCAGCGCA GATGCAGCCC 600
ATGCACCGCT ACGACGTGAG CGCCCTGCAG TACAACCTCA TGACCAGCTC GCAGACCTAC 660
ATGAACGGCT CGCCACCTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT 720
CTTGGCTCCA TGGGTTCCGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCC TGTGGTTACC 780
TCTTCCTCCC ACTCCAGGGC GCCCTGCCAG GCCGGGGACC TCCGGGACAT GATCAGCATG 840
TATCTCCCCG GCGCCGAGGT GCCGGAACCC GCCGCCCCCA GCAGACTTCA CATGTCCCAG 900
CACTACAGGA GCGGCGCGGT GCCCGCACG GCCATTAACG GCACACTGCC CCTCTCACAC 960
ATGTGAGGGC CGGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAAGAAAA ACGAGGGAAA 1020
TGGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAAACCCGG TACGCTCAA 1080
AAAAA

```

Seq ID NO: 347 Protein sequence  
Protein Accession #: CAA83435

80  
85

```

1 11 21 31 41 51
| | | | |
HSARMYNMME TELKPPGPQQ TSGGGGNGST AAAAGGNQKN SPDRVKRPMN AFMVWSRGQR 60
RKMAQENPKM HNSEISKRLG AEWKLLSETE KRPFIDEAKR LRALHMKHEP DYKYRPRRKT 120
KTLMKKDKYT LPPGLLAPGG NSMAGGVGVG AGLMAGVNRQ MDSYAHMNGW SNGSYSMQD 180
QLGYQPQHPGL NAHGAAQMOP MHRYDVSAHQ YNSMTSSQTY MNGSPTYSMS YSQQGTGMA 240
LGSMSGVVKS EASSPPVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPPEP AAPSR LHMSQ 300
HYQSGPVPPT AINGTLPLSH M

```



Seq ID NO: 348 DNA sequence  
Nucleotide Accession #: NM\_002638  
Coding sequence: 120-473

```
5      1      11      21      31      41      51
      |      |      |      |      |      |
      |      |      |      |      |      |
10     CAATACAGCT AAGGAATTAT CCCTTGTAAG TACCACAGAC CCGCCCTGGA GCCAGGCCAA 60
      GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
      TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCCTCAT CGCTGGGACG CTGGTTCTAG 180
      AGGCAGCTGT CACGGGAGTT CCTGTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240
      TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAAGTAA AGGTCAAGAT AAAGTCAAAG 300
      CGCAAGAGCC AGTCAAAGGT CCAGTCTCCA CTAAGCCTGG CTCTGCCCC ATTATCTTGA 360
      TCCGGTGCAG CATGTTGAAT CCCCCTAACC GCTGCTTGAA AGATACTGAC TGCCCAGGAA 420
15     TCAAGAAGTG CTGTGAAGG TCTTGCGGGA TGGCCTGTTT CGTCCCCAG TGAAGGGAGC 480
      CGGTCCTTGC TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540
      TGCTGCCCTT CCCCTTCCA CACTGTCCAT TCTTCTCCCT ATTCAGGATG CCCACGGCTG 600
      GAGCTGCCTC TCTCATCCAC TTTCCAATAA A
```

Seq ID NO: 349 Protein sequence:  
Protein Accession #: NP\_002629

```
25     1      11      21      31      41      51
      |      |      |      |      |      |
      |      |      |      |      |      |
      MRASSFLIVV VFLIAGTLVL EAAVTGVFVK GQDTVKGKRV FNGQDPVKGQ VSVKQDKVK 60
      AQEPVKGKPVV TKPGSCPIIL IRCAMLNPPN RCLKDTDCPG IKKCEGSCG MACFVPQ
```

Seq ID NO: 350 DNA sequence  
Nucleic Acid Accession #: NM\_007183  
Coding sequence: 75-2468

```
35     1      11      21      31      41      51
      |      |      |      |      |      |
      |      |      |      |      |      |
      GAATTCCGGA CAGGACGTGA AGATAGTTGG GTTTGGAGGC GGCCGCCAGG CCCAGGCCCG 60
      GTGGACCTGC CGCCATGCAG GACGGTAACT TCCTGCTGTC GGCCCTGCAG CCTGAGGCCG 120
      GCGTGTGCTC CTTGGCGCTG CCTCTGACC TGCAGCTGGA CCGCCGGGGC GCCGAGGGGC 180
      CGGAGGCCGA GCGGTGCGG GCAGCCCGCG TCCAGGAGCA GGTCCGCGCC CGCTCTTGC 240
      AGCTGGGACA GCAGCCGCGG CACAACGGGG CCGCTGAGCC CGAGCCTGAG GCCGAGACTG 300
40     CCAGAGGCAC ATCCAGGGGG CAGTACAGGC CCCTGCAGGC TGGCTTCAGC TCTCGCTCTC 360
      AGGGCCCTGAG TGGGGACAAG ACCTCGGGCT TCCGGCCCAT CGCCAAGCCG GCCTACAGCC 420
      CAGCCTCCTG GTCCTCCCGC TCCGCCGTGG ATCTGAGCTG CAGTCGGAGG CTGAGTTTCA 480
      CCCACAATGG GGGCAGCGCC TTTGGGGCCG CTGGGTACGG GGGTGGCCAG CCCACCCCTC 540
      CCATGCCCCA CAGGCCCGTG TCCTTCCATG AGCGCGGTGG GGTGGGGAGC CGGGCCGACT 600
45     ATGACACACT CTCCCTGCGC TCGCTGCGGG TGGGGCCCGG GGGCCTGGAC GACCGCTACA 660
      GCCTGGTGTG TGAGCAGCTG GAGCCCGCGG CCACCTCCAC CTACAGGGCC TTTGCGTACG 720
      AGCGCCAGGC CAGCTCCAGC TCCAGCCGGG CAGGGGGGCT GGAAGTGGCC GAGGCCACTG 780
      AGGTTTCCCC GAGCCGGACC ATCCGTGCCC CTGCCGTGCG GACCCTGCAG CGATTCCAGA 840
      GCAGCCACCG GAGCCGCGGG GTAGCGGGGG CAGTGCCGGG GGCCGTCTTG GAGCCAGTGG 900
50     CTCGAGCGCC ATCTGTGCGC AGCCTCAGCC TCAGCCTGGC TGAAGTGGGC CACCTGCCCG 960
      ACGTGCATGG GTTCAACAGC TACGGTAGCC ACCGAACCTT GCAGAGACTC AGCAGCGGTT 1020
      TTGATGACAT TGACCTGCCC TCAGCAGTCA AGTACCTCAT GGCTTCAGAC CCCAACCTGC 1080
      AGGTGCTGGG AGCGGCCTAC ATCCAGCACA AGTGCTACAG CGATGCAGCC GCCAAGAAGC 1140
      AGGCCCGCAG CCTTCAGGCC GTGCCTAGGC TGGTGAAGCT CTTCAACCAC GCCAACCAGG 1200
55     AAGTGCAGCG CCATGCCACA GGTGCCATGC GCAACCTCAT CTACGACAAC GCTGACAACA 1260
      AGCTGGCCCT GGTGGAGGAG AACGGGATCT TCGAGCTGCT GCGGACACTG CGGGAGCAGG 1320
      ATGATGAGCT TCGCAAAAAT GTCACAGGGA TCCTGTGGAA CCTTTCATCC AGCGACCACC 1380
      TGAAGGACCG CCTGGCCAGA GACACGCTGC AGCAGCTCAC GGACCTGGTG TTGAGCCCCC 1440
      TGTCTGGGGC TGGGGGTCCC CCCCTCATGG AGCAGAACGC CTCGGAGGCG GAGATCTTCT 1500
60     ACAACGCCAC CGGCTTCCTC AGGAACCTCA GCTCAGCCTC TCAGGCCACT CGCCAGAAGA 1560
      TGCGGGAGTG CCACGGGCTG GTGGACGCCC TGGTACCTC TATCAACCAC GCCCTGGACG 1620
      CGGGCAAATG CGAGGACAAG AGCGTGGAGA ACAGCGGTGT CGTCTGCGG AACCTGTCTT 1680
      ACCGCCTCTA CGAGGAGATG CCGCCGTCCG CGCTGCAGCG GCTGGAGGGT CGCGGCCGCA 1740
      GGGACCTGGC GGGGGCGCCG CCGGGAGAGG TCGTGGGCTG CTTACGCCC GAGAGCCGCG 1800
65     GGCTGCGCGA GCTGCCCTCT GCGGCGATG CGCTCACCTT CGCGAGGTG TCCAAGGACC 1860
      CCAAGGGCCT CGAGTGGCTG TGGAGCCCCC AGATCGTGGG GCTGTACAAC CGGCTGCTGC 1920
      AGCGCTGCGA GCTCAACCGG CACACGACGG AGGCGGCCGC CGGGGCGCTG CAGAACATCA 1980
      CGGCAGGCGA CCGCAGGTGG GCGGGGGTGC TGAGCCGCTT GGCCCTGGAG CAGGAGCGTA 2040
      TTCTGAACCC CCTGCTAGAC CGTGTGAGGA CCGCCGACCA CCACCAGCTG CGCTCACTGA 2100
70     CTGGCCTCAT CCGAAACCTG TCTCGGAACG CTAGGAACAA GGACGAGATG TCCACGAAGG 2160
      TGGTGAGCCA CCTGATCGAG AAGCTGCCAG CAGCGTGGG TGAGAAGTCG CCCCAGCCG 2220
      AGGTGCTGGT CAACATCATA GCTGTGCTCA ACAACCTGGT GGTGGCCAGC CCCATCGCTG 2280
      CCCGAGACCT GCTGTATTTT GACGGAATCC GAAAGCTCAT CTTTATCAAG AAGAAGCGGG 2340
75     ACAGCCCCGA CAGTGAGAAG TCCTCCCGGG CAGCATCCAG CCTCTGGGCC AACCTGTGGC 2400
      AGTACAACAA GCTCCACCGT GACTTTCCGG CGAAGGGCTA TCGGAAGGAG GACTTCCTGG 2460
      GCCCATAGGT GAAGCCTTCT GGAGGAGAAG GTGACGTGGC CCAGCGTCCA AGGGACAGAC 2520
      TCAGCTCCAG GCTGCTTGGC AGCCAGGCTG GGAGGAGAAG GCTAATGACG GAGGGGCCCC 2580
      TCGCTGGGGC CCTGTGTGTC ATCTTTGAGG GTCTTGGGCC ACCAGGAGGG GCAGGGTCTT 2640
80     ATAGCTGGGG ACTTGGCTTC CGCAGGGCAG GGGGTGGGGC AGGGCTCAAG GCTGCTCTGG 2700
      TGTATGGGGT GGTGACCCAG TCACATTGGC AGAGGTGGGG GTTGGCTGTG GCCTGGCAGT 2760
      ATCTTGGGAT AGCCAGCACT GGAATAAAG ATGGCCATGA ACAGTCACAA AAAAAAAAAA 2820
      AAAAGGAATT C
```

Seq ID NO: 351 Protein sequence  
Protein Accession #: NP\_009114.1

```
85     1      11      21      31      41      51
```



5  
10  
15

MQDGNFLLSA	LQPEAGVCSL	ALPSDLQLDR	RGAEGPEAER	LRAARVQEQV	RARLLQLGQQ	60
PRHNGAAEPE	PEAETARGTS	RGQYHTLQAG	FSSRSQGLSG	DKTSGFRPIA	KPAYSPASWS	120
SRSAVDLSCS	RRLSSAHNGG	SAFGAAGYGG	AQPTPPMPTR	PVSFHERGGV	GSRADYDTLS	180
LRSLRLGPGG	LDDRYSLVSE	QLEPAATSTY	RAFAYERQAS	SSSSRAGGLD	WPEATEVSPS	240
RTIRAPAVRT	LQRFQSSHRS	RGVCGAVPGA	VLEPVARAPS	VRSLSLSLAD	SGHLPDVHGF	300
NSYGSHTLQ	RLSSGFDDID	LPSAVKYLMA	SDPNLQVLGA	AYIQHKCYSD	AAAKKQARSL	360
QAVPRLVKLF	NHANQEVQRH	ATGAMRNLIY	DNADNKLALV	EENGIFELLR	TLREQDDELRL	420
KNVTGILWNL	SSSDHLKDRL	ARDTLEQLTD	LVLSPLSGAG	GPPLIQNAS	EAEIFYNATG	480
FLRNLSSASQ	ATRQKMRCH	GLVDALVTSI	NHALDAGKCE	DKSVENAVCV	LRNLSYRLYD	540
EMPPSALQRL	EGRGRDLAG	APPGEVVGCF	TPQSRRLREL	PLAADALTFA	EVSKDPKGLE	600
WLWSPQIVGL	YNRLQLRCEL	NRHTTEAAAAG	ALQNITAGDR	RWAGVLSRLA	LEQERILNPL	660
IDRVRTADHH	QLRSLTGLIR	NLSRNARNKD	EMSTKVVSHL	IEKLPGSVGE	KSPPAEVLVN	720
IIAVLNNLVV	ASPIAARDLL	YFDGLRKLIF	IKKKRSDSPDS	EKSSRAASSL	LANLWQYNKL	780
HRDFRAKGYR	KEDFLGP					

Seq ID NO: 352 DNA sequence  
Nucleic Acid Accession #: M31469  
Coding sequence: 1-651

20  
25  
30  
35

1	11	21	31	41	51	
ATGGCTGCGC	AGGGAGAGCC	CCAGGTCCAG	TTCAAACCTTG	TATTGGTTGG	TGATGGTGGT	60
ACTGGAAAAA	CGACCTTCGT	GAAACCTCAT	TTGACTGGTG	AATTTGAGAA	GAAGTATGTA	120
GCCACCTTGG	GTGTTGAGGT	TCATCCCCTA	GTGTTCCACA	CCAACAGAGG	ACCTATTAAG	180
TTCAATGTAT	GGGACACAGC	CGGCCAGGAG	AAATTCGGTG	GACTGAGAGA	TGGCTATTAT	240
ATCCAAGCCC	AGTGTGCCAT	CATAATGTTT	GATGTAACAT	CGAGAGTTAC	TTACAAGAAT	300
GTGCCTAACT	GGCATAGAGA	TCTGGTACGA	GTGTGTGAAA	ACATCCCCAT	TGTGTTGTGT	360
GGCAACAAAG	TGGATATTAA	GGACAGGAAA	GTGAAGGCGA	AATCCATTGT	CTTCCACCGA	420
AAGAAGAATC	TTCAGTACTA	CGACATTTCT	GCCAAAAGTA	ACTACAACCT	TGAAAAGCCC	480
TTCTCTGGC	TTGCTAGGAA	GCTCATTGGA	GACCCTAACT	TGGAATTTGT	TGCCATGCCT	540
GCTCTCGCCC	CACCAGAAGT	TGTCATGGAC	CCAGCTTTGG	CAGCACAGTA	TGAGCACGAC	600
TTAGAGGTTG	CTCAGACAAC	TGCTCTCCCG	GATGAGGATG	ATGACCTGTG	A	

Seq ID NO: 353 Protein sequence  
Protein Accession #: AAA36546

40  
45

1	11	21	31	41	51	
MAAQGEPPVQ	FKLVLVGDGG	TGKTTFVKRH	LTGEFEKKYV	ATLGVEVHPL	VFHTNRGPIK	60
FNWVDTAGQE	KFGGLRDGY	IQAQCAIMF	DVTSRVTYKN	VPNWHRLVR	VCENIPIVLC	120
GNKVDIKDRK	VKAKSIVFHR	KKNLQYVDIS	AKSNYNFEKP	FLWLARKLIG	DPNLEFVAMP	180
ALAPPEVMD	PALAAQYEH	LEVAQTALP	DEDDDL			

Seq ID NO: 354 DNA sequence  
Nucleic Acid Accession #: NM\_002820  
Coding sequence: 304-831

50  
55  
60  
65  
70  
75  
80

1	11	21	31	41	51	
CCGGTTCGCA	AAGAAGCTGA	CTTCAGAGGG	GGAAACTTTC	TTCTTTTAGG	AGGCGGTTAG	60
CCCTGTTCCA	CGAAGCCAGG	AGAACTGCTG	GCCAGATTAA	TTAGACATTG	CTATGGGAGA	120
CGTGTAACA	CACTACTTAT	CATTGATGCA	TATATAAAAC	CATTTTATTT	TCGCTATTAT	180
TTCAAGAGAA	GCGCCTCTGA	TTTGTTCCTT	TTTTCCCTTT	TTGCTCTTTC	TGGCTGTGTG	240
GTTTGAGAGAA	AGCACAGTTG	GAGTAGCCGG	TTGCTAAATA	AGTCCCGAGC	GCGAGCGGAG	300
ACGATGCAGC	GGAGACTGGT	TCAGCAGTGG	AGCGTCGCGG	TGTTCTTGCT	GAGCTACGCG	360
GTGCCCTCCT	GCGGGCGCTC	GGTGGAGGGT	CTCAGCCGCC	GCCTCAAAAG	AGCTGTGTCT	420
GAACATCAGC	TCCTCCATGA	CAAGGGGAAG	TCCATCCAAG	ATTTACGGCG	ACGATTCTTC	480
CTTCACCATC	TGATCGCAGA	AATCCACACA	GCTGAAATCA	GAGCTACCTC	GGAGGTGTCC	540
CCTAACTCCA	AGCCCTCTCC	CAACACAAAG	AACCACCCCG	TCCGATTTGG	GTCTGATGAT	600
GAGGGCAGAT	ACCTAACTCA	GGAAACTAAC	AAGGTGGAGA	CGTACAAAGA	GCAGCCGCTC	660
AAGACACCTG	GGAAGAAAAA	GAAAGGCAAG	CCCGGGAAAC	GCAAGGAGCA	GGAAAAGAAA	720
AAACGGCGAA	CTCGCTCTGC	TCTGTTAGAC	CTGGAGTGTA	CTGGGAGTGG	GCTAGAAGGG	780
GACCACTGT	CTGACACCTC	CACAACGTCG	CTGGAGCTCG	ATTCACGGTA	ACAGGCTTCT	840
CTGGCCCGTA	GCCTCAGCGG	GGTGCTCTCA	GCTGGGTTTT	GGAGCCTCCC	TTCTGCCTTG	900
GCTTGACAA	ACCTAGAATT	TTCTCCCTTT	ATGTATCTCT	ATCGATTGTG	TAGCAATTGA	960
CAGAGAATAA	CTCAGAATAT	TGTCTGCCTT	AAAGCAGTAC	CCCCCTACCA	CACACACCCC	1020
TGTCCTCCAG	CACCATAGAG	AGGCGCTAGA	GCCCATTCTT	CTTTCTCCAC	CGTCACCCAA	1080
CATCAATCCT	TTACCACTCT	ACCAAATAAT	TTCAATTTCA	AGCTTCAGAA	GCTAGTGACC	1140
ATCTTCATAA	TTTGCTGGAG	AAGTGTATTT	TTCCCTTTTA	CTCTCACACC	TGGGCAAACT	1200
TTCTTCAGTG	TTTTTCATTT	CTTACGTTCT	TTCACTTCAA	GGGAGAATAT	AGAAGCATTT	1260
GATATTATCT	ACAAACACTG	CAGAACAGCA	TCATGTCATA	AACGATTCTG	AGCCATTAC	1320
ACTTTTATTT	TAATTAAATG	TATTTAATTA	AATCTCAAAT	TTATTTTAAT	GTAAAGAACT	1380
TAAATTATGT	TTTAAACACA	TGCCTTAAAT	TTGTTTAAAT	AAATTTAACT	CTGGTTTCTA	1440
CCAGCTCATA	CAAAATAAAT	GGTTTCTGAA	AATGTTTAA	TATTAACCTA	CAAGGATATA	1500
GGTTTCTCTC	ATGTATCTTT	TTGTTTCAAT	GCAAGATGAA	ATAATTTTTC	TAGGGTAATG	1560
CCGTAGGAAA	AATAAACTT	CACATTTAAA	AAAAA			

Seq ID NO: 355 Protein sequence  
Protein Accession #: NM\_002820

85

1	11	21	31	41	51	
MQRRLVQQWS	VAVFLLSYAV	PSCGRSVEGL	SRRLKRAVSE	HQLLHDKGKS	IQDLRRRFFL	60
HHLIAEIHATA	EIRATSEVSP	NSKPSPTNKN	HPVRFSGSDE	GRYLTQETNK	VETYKEQPLK	120



TPGKKKKGKP GKRKEQKKK RRTRSAWLDS GVTGSGLEGD HLSDTSTTSL ELDSR

Seq ID NO: 356 DNA sequence  
Nucleic Acid Accession #: NM\_017522  
Coding sequence: 1-2100

5  
10  
15  
20  
25  
30  
35  
40  
45  
50

1	11	21	31	41	51	
ATGGGCCTCC	CCGAGCCGGG	CCCTCTCCGG	CTTCTGGCGC	TGCTGCTGCT	GCTGCTGCTG	60
CTGCTGCTGC	TGCGGCTCCA	GCATCTTGCG	GCGGCAGCGG	CTGATCCGCT	GCTCGGCGGC	120
CAAGGGCCGG	CCAAGGAGTG	CGAAAAGGAC	CAATTCCAGT	GCCGGAACGA	GCGCTGCATC	180
CCCTCTGTGT	GGAGATGCGA	CGAGGACGAT	GACTGCTTAG	ACCACAGCGA	CGAGGACGAC	240
TGCCCCAAGA	AGACCTGTGC	AGACAGTGAC	TTCACCTGTG	ACAACGGCCA	CTGCATCCAC	300
GAACGGTGGA	AGTGTGACGG	CGAGGAGGAG	TGTCCTGATG	GCTCCGATGA	GTCCGAGGCC	360
ACTTGCACCA	AGCAGGTGTG	TCCTGCAGAG	AAGCTGAGCT	GTGGACCCAC	CAGCCACAAG	420
TGTGTACCTG	CCTCGTGGCG	CTGCGACGGG	GAGAAGGACT	GCGAGGGTGG	AGCGGATGAG	480
GCCGGCTGTG	CTACCTCACT	GGGCACCTGC	CGTGGGGACG	AGTTCCAGTG	TGGGGATGGG	540
ACATGTGTCC	TTGCAATCAA	GCACTGCAAC	CAGGAGCAGG	ACTGTCCAGA	TGGGAGTGAT	600
GAAAGCTGGCT	GCCTACAGGG	GCTGAACGAG	TGTCTGCACA	ACAATGGCGG	CTGCTCACAC	660
ATCTGCACTG	ACCTCAAGAT	TGGCTTTGAA	TGCACGTGCC	CAGCAGGCTT	CCAGCTCCTG	720
GACCAGAAGA	CTTGTGGCGA	CATTGATGAG	TGCAAGGACC	CAGATGCCTG	CAGCCAGATC	780
TGTGTCAATT	ACTAAGGGCTA	TTTTAAGTGT	GAGTGTCTACC	CTGGCTGCGA	GATGGACCTA	840
CTGACCAAGA	ACTGCAAGGC	TGCTGCTGGC	AAGAGCCCAT	CCCTAATCTT	CACCAACCGC	900
ACGAGTGCGG	AGGATCGACC	TGTGAAGCGG	AACATATTCAC	GCCTCATCCC	CATGCTCAAG	960
AATGTCGTGG	CACTAGATGT	GGAAGTTGCC	ACCAATCGCA	TCTACTGGTG	TGACCTCTCC	1020
TACCGTAAGA	TCTATAGCGC	CTACATGGAC	AAGGCCAGTG	ACCCGAAAGA	GCGGGAGGTC	1080
CTCATTGACG	AGCAGTTGCA	CTCTCCAGAG	GGCCTGGCAG	TGGACTGGGT	CCACAAGCAC	1140
ATCTACTGGA	CTGACTCGGG	CAATAAGACC	ATCTCAGTGG	CCACAGTTGA	TGGTGGCCGC	1200
CGACGCACTC	TCTTCAGCCG	TAACCTCAGT	GAACCCCGGG	CCATCGCTGT	TGACCCCTCTG	1260
CGAGGGTTCA	TGTATTGGTC	TGACTGGGGG	GACCAGGCCA	AGATTGAGAA	ATCTGGGCTC	1320
AACGGTGTGG	ACCGGCAAAC	ACTGGTGTCA	GACAATATTG	AATGGCCCAA	CGGAATCACC	1380
CTGGATCTGC	TGAGCCAGCG	CTTGTACTGG	GTAGACTCCA	AGCTACACCA	ACTGTCCAGC	1440
ATTGACTTCA	GTGGAGGCAA	CAGAAAGACG	CTGATCTCCT	CCACTGACTT	CCTGAGCCAC	1500
CCTTTTGGGA	TAGCTGTGTT	TGAGGACAAG	GTGTTCTGGA	CAGACCTGGA	GAACGAGGCC	1560
ATTTTCAGTG	CAAATCGGCT	CAATGGCCTG	GAAATCTCCA	TCCTGGCTGA	GAACCTCAAC	1620
AACCCACATG	ACATTGTGAT	CTTCCATGAG	CTGAAGCAGC	CAAGAGCTCC	AGATGCCTGT	1680
GAGCTGAGTG	TCCAGCCTAA	TGGAGGCTGT	GAATACCTGT	GCCTTCCTGC	TCCTCAGATC	1740
TCCAGCCACT	CTCCCAAGTA	CACATGTGCC	TGTCCTGACA	CAATGTGGCT	GGGTCCAGAC	1800
ATGAAGAGGT	GCTACCGAGA	TGCAAATGAA	GACAGTAAGA	TGGGCTCAAC	AGTCACTGCC	1860
GCTGTTATCG	GGATCATCGT	GCCCATAGTG	GTGATAGCCC	TCCTGTGCAT	GAGTGGATAC	1920
CTGATCTGGA	GAAACTGGAA	GCGGAAGAAC	ACCAAAAGCA	TGAATTTTGA	CAACCCAGTC	1980
TACAGGAAAA	CAACAGAAGA	AGAAGATGAA	GATGAGCTCC	ATATAGGGAG	AACTGCTCAG	2040
ATTGGCCATG	TCTATCCTGC	ACGAGTGGCA	TTAAGCCTTG	AAGATGATGG	ACTACCCTGA	2100
GGATGGGATC	ACCCCTTCG	TGCCTCATGG	AATTCAGTCC	CATGCACTAC	ACTCCGGATG	2160
GTGTATGACT	GGATGAATGG	GTTTCTATAT	ATGGGTCTGT	GTGAGTGTAT	GTGTGTGTGT	2220
GATTTTTTTT	TTTAAATTTA	TGTTGCGGAA	AGGTAACCAC	AAAGTTATGA	TGAAGTGCAG	2280
ACATCCAAAG	GATGTGAGAG	TTTTTCTATG	TATAATGTTT	TATACACTTT	TTAACTGGTT	2340
GCACTACCCA	TGAGGAATTC	GTGGAATGGC	TACTGCTGAC	TAACATGATG	CACATAACCA	2400
AATGGGGGCC	AATGGCACAG	TACCTTACTC	ATCATTTAAA	AACTATATTT	ACAGAAGATG	2460
TTTGGTTGCT	GGGGGGCTTT	TTTAGGTTTT	GGGCATTTGT	TTTTTGTAAG	TAAGATGATT	2520
ATGCTTTGTG	GCTATCCATC	AACATAAGT				

Seq ID NO: 357 Protein sequence  
Protein Accession #: NP\_059992

55  
60  
65  
70

1	11	21	31	41	51	
MGLPEPGPLR	LLALLLLLLL	LLLLRLQHLA	AAAADPLLGG	QGPKECEKED	QFQCRNERCI	60
PSVWRCEDED	DCLDHSDEDD	CPKKTCAQSD	FTCDNGHCIH	ERWKCDGEEE	CPDGSDESEA	120
TCTKQVCPAE	KLSCGPSTSHK	CVPASWRCDG	EKDCEGGADE	AGCATSLGTC	RGDEFQCGDG	180
TCVLAIKHCN	QEQDCPDGSD	EAGCLQGLNE	CLHNNGGCSH	ICTDLKIGFE	CTCPAGFQLL	240
DQKTCGDIDE	CKDPDACSQI	CVNYKGYFKC	ECYPGCEMDL	LTKNCKAAAG	KSPSLIFTNR	300
TSAEDRPVKR	NYSRLIPMLK	NVVALDVEVA	TNRIYWCDLS	YRKIYSAYMD	KASDPKEREV	360
LIDEQLHSPE	GLAVDWWVHK	IYWTDSGNKT	ISVATVDGGR	RRTLFNRNLS	EPRAIADVPL	420
RGFMYSWDWG	DQAKIEKSLG	NGVDRQTLVS	DNIEWPNGIT	LDLLSQRLYW	VDSKLHQLSS	480
IDFSGGNRKT	LISSTDFLSH	PFGIIVFEDK	VFWTDLENEA	IFSANRLNGL	EISILAENLN	540
NPHDIVIFHE	LKQPRAPDAC	ELSVQPNGGC	EYLCPLAPQI	SSHSPKYTCA	CPDTMWLGPD	600
MKRCYRDANE	DSKMGSTVTA	AVIGIIVPIV	VIALLCMSGY	LIWRNWKRNK	TKSMNFDNPV	660
YRKTEEEDE	DELHIGRTAQ	IGHVYPARVA	LSLEDDGLP			

Seq ID NO: 358 DNA sequence  
Nucleic Acid Accession #: M27826  
Coding sequence: <1-503

75  
80  
85

1	11	21	31	41	51	
AGCCCAAGAA	ACATCTCACC	AATTTCAAAT	CTGATCTATT	CGGCTTAGCG	ACTGAAGATT	60
GACGCTGCCC	GATCGCCTCG	GAAGTCCCCT	GAACCATCAC	AGAAGCCGAG	CTTCGGGTAA	120
CTCTCACAGT	GGAGGGTAAG	TCCATCCCCT	GTTTAATCGA	TACGGGGGCT	ACCCACTCCA	180
CGTTGCCTTC	TTTTCAAGGG	CCTGTTTCCC	TGCCCCCAT	AACTGTTGTG	GGTATTGACG	240
GCCAAGCTTC	AAAACCCCTG	AAAACCTCCC	CACTCTGGTG	CCAACCTGGA	CAACACTCTT	300
TTATGCACTC	TTTTTTAGTT	ATCCCCACCT	GCCCCACTTC	CTTATTAGGC	CGAAATATTT	360
TAACCAAAAT	ATCTGCTTCC	CTGACTATTC	CTGGAGTACA	GCTACATCTC	ATTGCTGCCC	420
TTCTTCCCAA	TCCAAAGCCT	CCTTTGTGTC	CTCTAACATC	CCCACAATAT	CAGCCCTTAC	480
CACAAGACCT	CCCTTCAGCT	TAATCTCTCC	CACTCTAGGT	TCCCACGCCG	CCCCTAATCC	540
CACCTGAAGC	AGCCCTGAGA	AACATCGCCC	ATTCTCTCTC	CATACCACCC	CCCCAAAATT	600
TTGCGCGCTC	CAACACTTCA	ACACTATTTT	GTTTATTATT	TCTTATTAAT	ATCAGAAGGC	660



5  
 AGGAATGTCA GGCCTCTGAG CCCAGGCCAG GCCATCGCAT CCCCTGTGAC TTGCACGTAT 720  
 ACATCCAGAT GGCCTGAAGT AACTGAAGAT CCACAAAAGA AGTAAAAACA GCCTTAACTG 780  
 ATGACATTCC ACCATTGTGA TTTGTTCTCG CCCACCCCTA ACTGATCAAT GTACTTTGTA 840  
 ATCTCCCCCA CCCTTAAGAA GGTTCCTTTGT AATTCTCCCC ACCCTTGAGA ATGTACTTTG 900  
 TGAGATCCAC CCCTGCCAC CAGAGAACAA CCCCTTTGA TTGTAATTTT TTATTACCTT 960  
 CCCAAATCCT ATAAACAGC CCCACCCCTA TCTTCCTTCA CTGACTCTCT TTTCGGACTC 1020  
 AGCCACCGGC ACCCAGGTGA AATAAACAGC TTTATTGCTC AC

10  
 Seq ID NO: 359 Protein sequence  
 Protein Accession #: AAA65999

15  

1	11	21	31	41	51	
PKKHLTNFKS	DLFGLATEDW	RCPIASEVPW	TITEAELRVT	LTVEGKSIPC	LIDTGATHST	60
LPSFQGPVSL	APITVVGIDG	QASKPLKTPP	LWCQLGQHSF	MHSFLVIPTC	PLPLLGRNIL	120
TKLSASLTIP	GVQLHLIAL	LPNPKPPLCP	LTSPQYQPLP	QDLPSA		

20  
 Seq ID NO: 360 DNA sequence  
 Nucleic Acid Accession #: NM\_001854  
 Coding sequence: 162-5582

25  

1	11	21	31	41	51	
AACCATCAAA	TTTAGAAGAA	AAAGCCCTTT	GACTTTTTTC	CCCTCTCCCT	CCCCAATGGC	60
TGTGTAGCAA	ACATCCCTGG	CGATACCTTG	GAAAGGACGA	AGTTGGTCTG	CAGTCGCAAT	120
TTCGTGGGTT	GAGTTCACAG	TTGTGAGTGC	GGGGCTCGGA	GATGGAGCCG	TGGTCCTCTA	180
GGTGGAAGAA	GAAACGGTGG	CTCTGGGATT	TCACCGTAAC	AACCCTCGCA	TGACCTTCC	240
TCTTCCAAGC	TAGAGAGGTC	AGAGGAGCTG	CTCCAGTTGA	TGTACTAAAA	GCACTAGATT	300
TTCAACAATC	TCCAGAGGGA	ATATCAAAAA	CAACGGGATT	TTGCACAAAC	AGAAAGAATT	360
CTAAAGGCTC	AGATACTGCT	TACAGAGTTT	CAAAGCAAGC	ACAACCTCAGT	GCCCCAACAA	420
AACAGTTATT	TCCAGGTGGA	ACTTTCCTCC	AAGACTTTTC	AATACTATTT	ACAGTAAAAAC	480
CAAAAAAAGG	AATTCAGTCT	TTCCTTTTAT	CTATATATAA	TGAGCATGGT	ATTCAGCAAA	540
TTGGTGTGTA	GGTTGGGAGA	TCACCTGTTT	TTCTGTTTGA	AGACCACACT	GGAAAACCTG	600
CCCCAGAAGA	CTATCCCTTC	TTCAGAACTG	TTAACATCGC	TGACGGGAAG	TGGCATCGGG	660
TAGCAATCAG	CGTGGAGAAG	AAAACGTGTA	CAATGATTGT	TGATTGTAAG	AAGAAAACCA	720
CGAAACCACT	TGATAGAAGT	GAGAGAGCAA	TTGTTGATAC	CAATGGAATC	ACGGTTTTTG	780
GAACAAGGAT	TTTGGATGAA	GAAGTTTTTG	AGGGGGACAT	TCAGCAGTTT	TTGATCACAG	840
GTGATCCCAA	GGCAGCATAT	GACTACTGTG	AGCATATATAG	TCCAGACTGT	GACTCTTCAG	900
CACCCAAGGC	TGCTCAAGCT	CAGGAACCTC	AGATAGATGA	GTATGCACCA	GAGGATATAA	960
TCGAATATGA	CTATGAGTAT	GGGGAAGCAG	AGTATAAAGA	GGCTGAAAGT	GTAACAGAGG	1020
GACCCACTGT	AACTGAGGAG	ACAATAGCAC	AGACGGAGGC	AAACATCGTT	GATGATTTTC	1080
AAGAATACAA	CTATGGAACA	ATGGAAGT	ACCAGACAGA	AGCTCCTAGG	CATGTTTCTG	1140
GGACAAATGA	GCCAAATCCA	GTTGAAGAAA	TATTTACTGA	AGAATATCTA	ACGGGAGAGG	1200
ATTATGATTC	CCAGAGGAAA	AATTCGTAGG	ATACACTATA	TGAAAACAAA	GAAATAGACG	1260
GCAGGGATTG	TGATCTTCTG	GTAGATGGAG	ATTTAGGCGA	ATATGATTTT	TATGAATATA	1320
AAGAATATGA	AGATAAACCA	ACAAGCCCTC	CTAATGAAGA	ATTTGGTCCA	GGTGTACCAG	1380
CAGAAACTGA	TATTACAGAA	ACAAGCATAA	ATGGCCATGG	TGCATATGGA	GAGAAAGGAC	1440
AGAAAGGAGA	ACCAGCAGTG	GTTGAGCCTG	GTATGCTTGT	CGAAGGACCA	CCAGGACCAG	1500
CAGGACCTGC	AGGTATTATG	GGTCCTCCAG	GTCTACAAGG	CCCCACTGGA	CCCCCTGGTG	1560
ACCCTGGCGA	TAGGGGCCCT	CCAGGACGTC	CTGGCTTACC	AGGGGCTGAT	GGTCTACCTG	1620
GTCTCTCTGG	TACTATGTTG	ATGTTACCGT	TCCGTTATGG	TGGTGTATGG	TCCAAAGGAC	1680
CAACCATCTC	TGCTCAGGAA	GCTCAGGCTC	AAGCTATTCT	TCAGCAGGCT	CGGATTGCTC	1740
TGAGAGGCC	ACCTGGCCCA	ATGGGTCTAA	CTGGAAGACC	AGGTCTGTG	GGGGGGCCTG	1800
GTTTATCTGG	GGCCAAAGGT	GAGAGTGGTG	ATCCAGGTCC	TCAGGGCCCT	CGAGGCGTCC	1860
AGGGTCCCC	TGGTCCAACG	GGAAAACCTG	GAAAAAGGGG	TCGTCCAGGT	GCAGATGGAG	1920
GAAGAGGAAT	GCCAGGAGAA	CCTGGGGCAA	AGGGAGATCG	AGGGTTTGAT	GGACTTCCGG	1980
GTCTGCCAGG	TGACAAAGGT	CACAGGGGTG	AACGAGGTCC	TCAAGTCCCT	CCAGGTCCCTC	2040
CTGGTGTATG	TGGAATGAGG	GGAGAAGATG	GAGAAATTGG	ACCAAGAGGT	CTTCCAGGTG	2100
AAGCTGGCCC	ACGAGGTTTG	CTGGGTCCAA	GGGGAACTCC	AGGAGCTCCA	GGGCAGCCTG	2160
GTATGGCAGG	TGTAGATGGC	CCCCCAGGAC	CAAAAGGGAA	CATGGGTCCC	CAAGGGGAGC	2220
CTGGGCCTCC	AGGTCAACAA	GGGAATCCAG	GACCTCAGGG	TCTTCTGGT	CCACAAGGTC	2280
CAATTGGTCC	TCCTGGTGAA	AAAGGACCAC	AAGGAAAACC	AGGACTTGCT	GGACTTCCTG	2340
GTGCTGATGG	GCCTCCTGGT	CATCCTGGGA	AAGAAGGCCA	GTCTGGAGAA	AAGGGGGCTC	2400
TGGGTCCCC	TGGTCCACAA	GGTCCATTG	GATNNCCGGG	CCCCCGGGGA	GTAAGGGGAG	2460
CAGATGGTGT	CAGAGGTCTC	AAGGGATCTA	AAGGTGAAAA	GGGTGAAGAT	GGTTTTCCAG	2520
GATTCAAAGG	TGACATGGGT	CTAAAAGGTG	ACAGAGGAGA	AGTTGGTCAA	ATTGGCCCAA	2580
GAGGGNAAGA	TGGCCCTGAA	GGACCCAAAG	GTGAGCAGG	CCCAACTGGA	GACCCAGGTC	2640
CTTCAGGTCA	AGCAGGAGAA	AAGGGAAAAC	TTGGAGTTCC	AGGATTACCA	GGATATCCAG	2700
GAAGACAAGG	TCCAAAGGGT	TCCACTGGAT	TCCCTGGGTT	TCCAGGTGCC	AATGGAGAGA	2760
AAGGTGCACG	GGGAGTAGCT	GGCAAACCAG	GCCCTCGGGG	TCAGCGTGGT	CCAACGGGTC	2820
CTCGAGGTTT	AAGAGGTGCA	AGAGGTCCCA	CTGGGAAAAC	TGGGCCAAAG	GGCACTTCAG	2880
GTGGCGATGG	CCCTCCTGGC	CCTCCAGGTG	AAAGAGGTCC	TCAAGGACCT	CAGGGTCCAG	2940
TTGGATTCCC	TGGACCAAAA	GGCCCTCCTG	GACCACCAGG	AAGGATGGGC	TGCCCAGGAC	3000
ACCCTGGGCA	ACGTGGGGAG	ACTGGATTTC	AAGGCAAGAC	CGGCCCTCCT	GGGCCAGGGG	3060
GAGTGGTTGG	ACCACAGGGA	CCAACCGGTG	AGACTGGTCC	AATAGGGGAA	CGTGGGTATC	3120
CTGGTCTCTC	TGGCCCTCCT	GGTGAGCAAG	GTCTTCTTGG	TGCTGCAGGA	AAAGAAGGTG	3180
CAAAGGGTGA	TCCAGGTCTT	CAAGGTATCT	CAGGGAAGA	TGGACCAGCA	GGATTACGTG	3240
GTTTCCAGG	GGAAAGAGGT	CTTCTCTGGG	CTCAGGGTGC	ACCTGGACTG	AAAGGAGGGG	3300
AAGGTCCCCA	GGGCCACCA	GGTCCAGTTG	GCTCACCAGG	AGAACGTGGG	TCAGCAGGTA	3360
CAGCTGGCCC	AATTGGTTTA	CGAGGGCGCC	CGGGACCTCA	GGGTCTCTCT	GGTCCAGCTG	3420
GAGAGAAAGG	TGCTCCTGGA	GAAAAAGGTC	CCCAAGGGCC	TGCAGGGAGA	GATGGAGTTC	3480
AAGGTCTCTG	TGGTCTCCCA	GGGCCAGCTG	GTCTGCGCGG	CTCCCTGGG	GAAGACGGAG	3540
ACAAGGGTGA	AATTGGTGAG	CCGGGACAAA	AAGGCACAA	GGGTGGCAAG	GGAGAAAAATG	3600
GCCCTCCCGG	TCCCCCAGGT	CTTCAACGAC	CAGTTGGTGC	CCCTGGAATT	GCTGGAGGTG	3660
ATGGTGAACC	AGGTCTTAGA	GGACAGCAGG	GGATGTTTGG	GCAAAAAGGT	GATGAGGGTG	3720
CCAGAGGCTT	CCCTGGACCT	CCTGGTCCAA	TAGGTCTTCA	GGGTCTGCCA	GGCCACCTG	3780
GTGAAAAAGG	TGAAAATGGG	GATGTTGGTC	CATGGGGGCC	ACCTGGTCTT	CCAGGCCCAA	3840



	GAGGCCCTCA	AGGTCCCAAT	GGAGCTGATG	GACCACAAGG	ACCCCCAGGT	TCTGTTGGTT	3900
	CAGTTGGTGG	TGTTGGAGAA	AAGGGTGAAC	CTGGAGAAGC	AGGAAACCCA	GGGCCTCCTG	3960
	GGGAAGCAGG	TGTAGGCGGT	CCCAAAGGAG	AAAAGAGGAG	GAAAGGGGAA	GCTGGTCCAC	4020
5	CTGGAGCTGC	TGGACCTCCA	GGTGCCAAGG	GGCCGCCAGG	TGATGATGGC	CCTAAGGGTA	4080
	ACCCGGGTCC	TGTTGGTTTT	CCTGGAGATC	CTGGTCCTCC	TGGGGAACCT	GGCCCTGCAG	4140
	GTCAAGATGG	TGTTGGTGGT	GACAAGGGTG	AAGATGGAGA	TCCTGGTCAA	CCGGGTCTCT	4200
	CTGGCCCATC	TGGTGAGGCT	GGCCCACCAG	GTCCTCCTGG	AAAACGAGGT	CCTCCTGGAG	4260
	CTGCAGGTGC	AGAGGGAAGA	CAAGGTGAAA	AAGGTGCTAA	GGGGGAAGCA	GGTGACAGAAG	4320
10	GTCTCTCTGG	AAAAACCGGC	CCAGTCGGTC	CTCAGGGACC	TGCAGGAAAG	CCTGCTCCAG	4380
	AAGGTCTTCG	GGGCATCCCT	GGTCCTGTGG	GAGAACAAGG	TCTCCCTGGA	GCTGCAGGCC	4440
	AAGATGGACC	ACCTGGTCCT	ATGGGACCTC	CTGGCTTACC	TGGTCTCAAA	GGTGACCCTG	4500
	GCTCCAAGGG	TGAAAAGGGA	CATCCTGGTT	TAATTGGCCT	GATTGGTCCT	CCAGGAGAAC	4560
	AAGGGGAAAA	AGGTGACCGA	GGGCTCCCTG	GAATCAAGG	ATCTCCAGGA	GCAAAAGGGG	4620
15	ATGGGGGAAT	TCCTGGTCCT	GCTGGTCCCT	TAGGTCCACC	TGGTCTCTCA	GGCTTACCAG	4680
	GTCTCAAGG	CCCAAAGGGT	AACAAAGGCT	CTACTGGACC	CGCTGGCCAG	AAAGGTGACA	4740
	GTGGTCTTCC	AGGGCCTCCT	GGGCCTCCAG	GTCCACCTGG	TGAAGTCATT	CAGCCTTTAC	4800
	CAATCTTGTC	CTCCAAAAAA	ACGAGAAGAC	ATACTGAAGG	CATGCAAGCA	GATGCAGATG	4860
	ATAATATTCT	TGATTACTCG	GATGGAATGG	AAGAAATATT	TGGTTCCTCT	AATTCCTCTG	4920
20	AACAAGACAT	CGAGCATATG	AAATTTCCAA	TGGGTACTCA	GACCAATCCA	GCCCCAACTT	4980
	GTAAAGACCT	GCAACTCAGC	CATCCTGACT	TCCCAGATGG	TGAATATTGG	ATTGATCCTA	5040
	ACCAAGGTTG	CTCAGGAGAT	TCCTTCAAAG	TTTACTGTAA	TTTCACATCT	GGTGGTGAGA	5100
	CTTGCAATTA	TCCAGACAAA	AAATCTGAGG	GAGTAAGAA	TTCATCATGG	CCAAAGGAGA	5160
	AACCAGGAAG	TTGGTTTAGT	GAATTTAAGA	GGGGAAAACT	GCTTTCATAC	TTAGATGTTG	5220
25	AAGGAAATTC	CATCAATATG	GTGCAAATGA	CATTCTGAA	ACTTCTGACT	GCCTCTGCTC	5280
	GGCAAAATTT	CACCTACCAC	TGTCATCAGT	CAGCAGCCTG	GTATGATGTG	TCATCAGGAA	5340
	GTTATGACAA	AGCACTTCGC	TTCCTGGGAT	CAATGATGA	GGAGATGTCC	TATGACAATA	5400
	ATCCTTTTAT	CAAAACACTG	TATGATGGTT	GTACGTCCAG	AAAAGGCTAT	GAAAAAACTG	5460
	TCATTGAAAT	CAATACACCA	AAATACCTAT	AAGTACCTAT	TGTTGATGTC	ATGATCAGTG	5520
30	ACTTTGGTGA	TCAGAATCAG	AAGTTCGGAT	TTGAAGTTGG	TCCTGTTTGT	TTTCTTGGCT	5580
	AAGATTAAGA	CAAAGAACAT	ATCAAATCAA	CAGAAAATGT	ACCTTGGTGC	CACCAACCCA	5640
	TTTTGTGCCA	CATGCAAGTT	TTGAATAAGG	ATGTATGGAA	AACAACGCTG	CATATACAGG	5700
	TACCATTTAG	GAAATACCGA	TGCCTTTGTG	GGGGCAGAA	CACAGACAAA	AGCTTTGAAA	5760
	ATCATAAAGA	TATAAGTTGG	TGTGGCTAAG	ATGGAAACAG	GGCTGATTCT	TGATTCCTCA	5820
35	TTCTCAACTC	TCCTTTTCCT	ATTTGAATTT	CTTTGGTGCT	GTAGAAAACA	AAAAAAGAAA	5880
	AATATATATT	CATAAAAAAT	ATGGTGCTCA	TTCTCATCCA	TCCAGGATGT	ACTAAAACAG	5940
	TGTGTTTAA	AAATTGTAAT	TATTTTGTGT	ACAGTTCTAT	ACTGTTATCT	GTGTCCATTT	6000
	CCAAACTTGG	CACGTGTCCC	TGAATTCCGC	TGACTCTAAT	TTATGAGGAT	GCCGAACCTC	6060
40	GATGGCAATA	ATATATGTAT	TATGAAAATG	AAGTTATGAT	TTCCGATGAC	CCTAAGTCCC	6120
	TTTCTTTGGT	TAATGATGAA	ATTCTTTTGT	GTGTGTTT			

Seq ID NO: 361 Protein sequence  
Protein Accession #: NP\_001845

45	1	11	21	31	41	51	
	MEPWSSRWKT	KRWLWDFTVT	TLALTFLFQA	REVRGAAPVD	VLKALDFHNS	PEGISKTTGF	60
	CTNRKNSKGS	DTAYRVSKQA	QLSAPTQQLF	PGGTFFPEDFS	ILFTVKPKKG	IQSFLLSIYN	120
	EHGIQQIGVE	VGRSPVFLFE	DHTGKPAPED	YPLFRTVNIA	DGKWHRVAIS	VEKKTVTMIV	180
50	DCKKKTTKPL	DRSERAIVDT	NGITVFGTRI	LDEEVFEGDI	QQFLITGDPK	AAYDYCEHYS	240
	PDCDSSAPKA	AQAQEPQIDE	YAPEDIIEYD	YEYGEAEYKE	AESVTEGPTV	TEETIAQTEA	300
	NIVDDFQEYN	YGTMESYQTE	APRHVSGTNE	PNPVEEIFTE	EYLTGEDYDS	QRKNSIEDTLY	360
	ENKEIDGRDS	DLVVDGDLGE	YDFYBYKEYE	DKPTSPPNNE	FGPGVPAETD	ITETSINGHG	420
	AYGEKGQKGE	PAVVEPGMLV	EGPPGPAGPA	GIMGPPGLQG	PTGPPGDPGD	RGPPGRPLGP	480
55	GADGLPGPPG	TMLMLPFRYG	GDGSKGPTIS	AQEAQAQAIL	QQARIALRGP	PGPMGLTGRP	540
	GPVGGPGSSG	AKGESGDPGP	QGPRGVQGGP	GPTGKPGKRG	RPGADGGRGM	PGEPAKGRDR	600
	GFDGLPLPLG	DKGHRGERGP	QGPPGPPGDD	GMRGEDGEIG	PRGLPGEAGP	RGLLGRPGTP	660
	GAPGQPGMAG	VDGPPGPKGN	MGPQGEPPGP	GQQGNPQPQG	LPGPQGPPIG	PGEKGPQKGP	720
	GLAGLPGADG	PPGHPGKEGQ	SGEKGALGPP	GPQGPPIGXP	PRGVKGADGV	RGLKGSKEGK	780
60	GEDGFPGFKG	DMGLKGDRGE	VQIGPRGXD	GPEGPKGRAG	PTGDPGPSGQ	AGEKGKLGVP	840
	GLPGYPGRQG	PKGSTGFPGF	PGANGEKGAR	GVAGKPGPRG	QRGPTGPRGS	RGARGPTGKP	900
	GPKGTSBGDG	PPGPPGERGP	QGPPGVPVGF	GPKGPPGPPG	RMGCPGHPGQ	RGETGFQGKT	960
	GPPGPGGVVG	PQGPTGETGP	IGERGPYPGP	GPPGEQGLPG	AAGKEGAKGD	PGPQGISGKD	1020
	GPAGLRGFPG	ERGLPGAQGA	PGLKGGEGPQ	GPPGPVGSPP	ERGSAGTAGP	IGLRGRPGPQ	1080
65	GPPGPAGEKG	APGEKGPQGP	AGRDGVQGPV	GLPGPAGPAG	SPGEDGDKGE	IGEPGQKGSK	1140
	GGKGENGPPG	PPGLQGPVGA	PGIAGDGDEP	GPRGQGMFG	QKGDGARGF	PGPPGPGLGQ	1200
	GLPGPPGEKG	ENGDVGPWGP	PGPPGPRGPQ	GPNGADGPQG	PPGSVGSVGG	VGEKGEPGEA	1260
	GNPGPPGEAG	VGGPKGERGE	KGEAGPPGAA	GPPGAKGPPG	DDGPKGNPGP	VGFPDGPDPG	1320
	GELGPAGQDG	VGGDKGEDGD	PGQPPGPPGS	GEAGPPGPPG	KRGPPGAAGA	EGRQGEKGAK	1380
70	GEAGAEGPPG	KTGPVGPQGP	AGKPGPEGLR	GIPGPVGEQG	LPGAAGQDGP	PGPMGPPPLP	1440
	GLKGDPSGSK	EKGHPGLIGL	IGPPGEQGEK	GDRGLPGTQG	SPGAKGDGGI	PGPAGPLGPP	1500
	GPPGLPGPQG	PKGNKGSTGP	AGQKGDGSLP	GPPGPPGPPG	EVIQPLPLLS	SKKTRRHTEG	1560
	MQADADDNII	DYSDGMEEIF	GSLNSLKQDI	EHMKFPMGTQ	TNPARTCKDL	QLSHPDFPDG	1620
	EYWIDPNQGC	SGDSFKVYCN	FTSGGETCIY	PDKKSEGVRI	SSWPKEKPGS	WFSEFKRGKL	1680
75	LSYLDVEGNS	INMVQMTFLK	LLTASARQNF	TYHCHQSAAW	YDVSSGSYDK	ALRFLGSNDE	1740
	EMSYDNNPFI	KTLYDGCTSR	KGYEKTIVIE	NTPKIDQVPI	VDVMISDFGD	QNQKFGFEVG	1800
	PVCFLG						

Seq ID NO: 362 DNA sequence  
Nucleic Acid Accession #: NM\_003107  
Coding sequence: 351-1775

85	1	11	21	31	41	51	
	TTCCCCAGCA	TTGAGAAAC	TCCTCTCTAC	TTTAGCACGG	TCTCCAGACT	CAGCCGAGAG	60
	ACAGCAAAC	GCAGCGCGGT	GAGAGAGCGA	GAGAGAGGGA	GAGAGAGACT	CTCCAGCCTG	120
	GGAACATAA	CTCCTCTGCG	AGAGGCGGAG	AACTCCTTCC	CCAAATCTTT	TGGGGACTTT	180



	TCTCTCTTTA	CCCACCTCCG	CCCCTGCGAG	GAGTTGAGGG	GCCAGTTCGG	CCGCCGCGCG	240
	CGTCTTCCCG	TTCGGCGTGT	GCTTGGCCCG	GGGAACCGGG	AGGGCCCGGC	GATCGCGCGG	300
	CGGCCGCGCG	GAGGGTGTGA	GCGCGCGTGG	GCGCCCGCCG	AGCCGAGGCC	ATGGTGACAGC	360
5	AAACCAACAA	TGCCGAGAAC	ACGGAAGCGC	TGCTGGCCCG	CGAGAGCTCG	GACTCGGGCG	420
	CCGGCCTCGA	GCTGGGAATC	GCCTCTCTCC	CCACGCCCGG	CTCCACCGCC	TCCACGGGCG	480
	GCAAGGCCGA	CGACCCGAGC	TGGTGCAAGA	CCCCGAGTGG	GCACATCAAG	CGACCCATGA	540
	ACGCCTTCAT	GGTGTGGTCG	CAGATCGAGC	GGCGCAAGAT	CATGGAGCAG	TCGCCCAGCA	600
	TGCACAACGC	CGAGATCTCC	AAGCGGCTGG	GCAAACGCTG	GAAGCTGCTC	AAAGACAGCG	660
10	ACAAGATCCC	TTTCATTCTGA	GAGGCGGAGC	GGCTGCGCCT	CAAGCACATG	GCTGACTACC	720
	CCGACTACAA	GTACCGGCC	AGGAAGAAGG	TGAAGTCCGG	CAACGCCAAC	TCCAGCTCCT	780
	CGGCCGCGCG	CTCCTCCAAG	CCGGGGGAGA	AGGGAGACAA	GGTCGGTGGC	AGTGGCGGGG	840
	GCGGCCATGG	GGGCGGCGGC	GGCGGCGGGA	GCAGCAACGC	GGGGGGAGGA	GGCGGCGGTG	900
	CGAGTGGCGG	CGGCGCCAAC	TCCAAACCGG	CGCAGAAAAA	GAGCTGCGGC	TCCAAAGTGG	960
	CGGGCGGCGC	GGGCGGTGGG	GTTAGCAAA	CGCACGCCAA	GCTCATCTTG	GCAGGCGGCG	1020
15	GCGGCGGCGG	GAAAGCAGCG	GCTGCCGCGG	CCGCTCTCTT	CGCCGCCGAA	CAGGCGGGGG	1080
	CCGCCGCCCT	GCTGCCCTTG	GGCGGCGCGG	CCGACCACCA	CTCGCTGTAC	AAGGCGCGGA	1140
	CTCCAGCGC	CTCGGCTTCC	GCCTCTCTCG	CAGCCTCGGC	CTCCGACGCG	CTCGCGGCCC	1200
	CGGGCAAGCA	CCTGGCGGAG	AAGAAGTGTA	AGCGCGTCTA	CCTGTTCCGC	GGCCTGGGCA	1260
20	CGTCGTCGTC	GCCCGTGGGC	GGCGTGGGCG	CGGGAGCCGA	CCCCAGCGAC	CCCCTGGGCC	1320
	TGTACGAGGA	GGAGGGCGCG	GGCTGCTCGC	CCGACGCGCC	CAGCCTGAGC	GGCCGACGCA	1380
	GCGCCGCCCT	GTCCCCCGCC	GCCGCGCGCT	CGCCCGCCGA	CCACCGCGGC	TACGCCAGCC	1440
	TGCGCGCGCG	CTCGCCCGCC	CCGTCCAGCG	CGCCCTCGCA	CGCGTCTCTC	TCGGCCTCGT	1500
	CCCACTCTCT	CTCTTCTCTC	TCCTCGGGCT	CCTCGTCTCT	CGACGACGAG	TTCAAGACG	1560
25	ACCTGCTCGA	CCTGAACCCC	AGCTCAAAC	TTGAGAGCAT	GTCCCTGGGC	AGCTTCAGTT	1620
	CGTCGTCGGC	GCTCGACCGG	GACCTGGATT	TTAACTTCGA	GCCCGGCTCC	GGCTCGCACT	1680
	TCGAGTTCCC	GGACTACTGC	ACGCCCGAGG	TGAGCGAGAT	GATCTCGGGA	GACTGGCTCG	1740
	AGTCCAGCAT	CTCCAACCTG	GTTTTACCT	ACTGAAGGGC	GCGCAGGCAG	GGAGAAGGGC	1800
	CGGGGGGGGT	AGGAGAGGAG	AAAAAAAAG	TGAAAAAAG	AAACGAAAAG	GACAGACGAA	1860
30	GAGTTTAAAG	AGAAAAGGGA	AAAAAGAAAG	AAAAAGTAAG	CAGGGCTCGT	TCGCCCCGCT	1920
	TCTCGTCGTC	GGATCAAGGA	GCGCGGCGGC	GTTTTGGACC	CGCGCTCCCA	TCCCCACCT	1980
	TCCCCGGCCG	GGGACCCACT	CTGCCCAGCC	GGAGGGACGC	GGAGGAGGAA	GAGGGTAGAC	2040
	AGGGGCGACC	TGTGATTGTT	GTTATTGATG	TTGTTGTTGA	TGGCAAAAAA	AAAAAGCGAC	2100
	TTGAGTTTGG	CTCCCCCTTG	CTTGAAGAGA	CCCCCTCCCC	CTTCCAACGA	GCTTCCGGAC	2160
35	TTGTCTGCAC	CCCCAGCAAG	AAGGCGAGTT	AGTTTTCTAG	AGACTTGAAG	GAGTCTCCCC	2220
	CTTCTGTCAT	CACCACCTTG	GTTTTGTTTT	ATTTTGCTTC	TTGGTCAAGA	AAGGAGGGGA	2280
	GAACCCAGCG	CACCCCTCCC	CCCCTTTTTT	TAAACGCGTG	ATGAAGACAG	AAGGCTCCGG	2340
	GGTGACGAAT	TTGGCCGATG	GCAGATGTTT	TGGGGGAACG	CCGGGACTGA	GAGACTCCAC	2400
	GCAGGCGAAT	TCCCGTTTGG	GGCCTTTTTT	TCCTCCCTCT	TTTCCCCTTG	CCCCCTCTGC	2460
40	AGCCGGAGGA	GGAGATGTTG	AGGGGAGGAG	GCCAGCCAGT	GTGACCGGCG	CTAGGAAATG	2520
	ACCCGAGAAC	CCCGTTGGAA	GCGCAGCAGC	GCGCAGTAGG	GGCGGGGGCG	GAGGAGGACA	2580
	CGAACTGGAA	GGGGGTTTAC	GGTCAAAC	AAATGGATTT	GCACGTTGGG	GAGCTGGCGG	2640
	CGGCGGCTGC	TGGGCCTCCG	CCTTCTTTTC	TACGTGAAAT	CAGTGAGGTG	AGACTTCCCA	2700
45	GACCCCGGAG	GCGTGGAGGA	GAGGAGACTG	TTTGATGTGG	TACAGGGGCA	GTCACTGGAG	2760
	GGCGAGTGGT	TTCGGAAAAA	AAAAAGAAA	AAAAGGG			

Seq ID NO: 363 Protein sequence  
Protein Accession #: NP\_003098

50	1	11	21	31	41	51	
	MVQQTNNNAEN	TEALLAGESS	DSGAGLELGI	ASSPTPGSTA	STGGKADDP	WCKTPSGHIK	60
	RPMNAFMVWS	QIERRKIMEQ	SPDMHNAEIS	KRLGKRWKLL	KDSKIPFIR	EAERLRLKHM	120
	ADYPDYKYRP	RKKVKSNGAN	SSSSAAASSK	PGEKGDVKG	SGGGGHGGGG	GGGSSNAGGG	180
55	GGGASGGGAN	SKPAQKKSCG	SKVAGGAGGG	VSKPHAKLIL	AGGGGGGKAA	AAAAASF	240
	QAGAAALLPL	GAAADHHS	KARTPSASAS	ASSAASASAA	LAAPGKHLAE	KKVKRVYLF	300
	GLGTSSSPVG	GVGAGADPSD	PLGLYEEEGA	GCSPDAPSL	GRSSAASSPA	AGRSPADHRG	360
	YASLRAASPA	PSSAPSHASS	SASSHSSSSS	SSGSSSSDDE	FEDDLLDLNP	SSNFESMSLG	420
	SFSSSSALDR	DLDFNFEPGS	GSHFEFPDYC	TPEVSEMISG	DWLESSISNL	VFTY	

Seq ID NO: 364 DNA sequence  
Nucleic Acid Accession #: U10860  
Coding sequence: 123-2204

65	1	11	21	31	41	51	
	TGCCGGCTGC	TCCTCGACCA	GGCCTCCTTC	TCAACCTCAG	CCCGCGGCGC	CGACCCCTTC	60
	GGCACCTTCC	CGCCCCGTCT	CGTACTGTCT	CGGTCAACGC	CGCGGCTCCG	GCCCTGGCCC	120
	CGATGGCTCT	GTGCAACGGA	GACTCCAAGC	TGGAGAATGC	TGGAGGAGAC	CTTAAGGATG	180
70	GCCACCACCA	CTATGAAGGA	GCTGTTGTCA	TTCTGGATGC	TGGTGCTCAG	TACGGGAAAG	240
	TCATAGACCG	AAGAGTGAGG	GAAGTGTTCG	TGCAGTCTGA	AATTTTCCCC	TTGGAAACAC	300
	CAGCATTTGC	TATAAAGGAA	CAAGGATTCC	GTGCTATTAT	CATCTCTGGA	GGACCTAATT	360
	CTGTGTATGC	TGAAGATGCT	CCCTGGTTTG	ATCCAGCAAT	ATTCACTATT	GGCAAGCCTG	420
	TTCTTGGAAT	TTGCTATGGT	ATGCAGATGA	TGAATAAGGT	ATTTGGAGGT	ACTGTGCACA	480
75	AAAAAAGTGT	CAGAGAAGAT	GGAGTTTTCA	ACATTAGTGT	GGATAATACA	TGTTCAATTAT	540
	TCAGGGGCCT	TCAGAAGGAA	GAAGTTGTTT	TGCTTACACA	TGGAGATAGT	GTAACAAAG	600
	TAGCTGATGG	ATTCAAGGTT	GTGGCACGTT	CTGGAAACAT	AGTAGCAGGC	ATAGCAAATG	660
	AATCTAAAAA	GTTATATGGA	GCACAGTTCC	ACCCTGAAGT	TGGCCTTACA	GAAAATGGAA	720
	AAGTAATACT	GAAGAATTTT	CTTTATGATA	TAGCTGGATG	CAGTGGAACC	TTCACCGTGC	780
80	AGAACAGAGA	ACTTGAGTGT	ATTCGAGAGA	TCAAAGAGAG	AGTAGGCACG	TCAAAGTTT	840
	TGGTTTACT	CAGTGGTGGA	GTAGACTCAA	CAGTTTGATC	AGCTTTGCTA	AATCGTGCTT	900
	TGAACCAAGA	ACAAGTCATT	GCTGTGCACA	TTGATAATGG	CTTTATGAGA	AAACGAGAAA	960
	GCCAGTCTGT	TGAAGAGGCC	CTCAAAAAGC	TTGGAATTCA	GGTCAAAGTG	ATAAATGCTG	1020
	CTCATTTCTT	CTACAATGGA	ACAACAACCC	TACCAATATC	AGATGAAGAT	AGAACCCAC	1080
85	GGAAAAAGAT	TAGCAAAACG	TTAAATATGA	CCACAAGTCC	TGAAGAGAAA	AGAAAAATCA	1140
	TTGGGGATAC	TTTTGTAAAG	ATTGCCAATG	AAGTAATTGG	AGAAATGAAC	TTGAAACCAG	1200
	AGGAGGTTTT	CCTTGCCCAA	GGTACTTTAC	GGCTGATCT	AATTGAAAGT	GCATCCCTTG	1260



TTGCAAGTGG CAAAGCTGAA CTCATCAAAA CCCATCACAA TGACACAGAG CTCATCAGAA 1320  
 AGTTGAGAGA GGAGGGAAAA GTAATAGAAC CTCTGAAAGA TTTTCATAAA GATGAAGTGA 1380  
 GAATTTTGGG CAGAGAACTT GGACTTCCAG AAGAGTTAGT TTCCAGGCAT CCATTTCCAG 1440  
 5 GTCTTGGCCT GGCAATCAGA GTAATATGTG CTGAAGAACC TTATATTTGT AAGGACTTTC 1500  
 CTGAAACCAA CAATATTTTG AAAATAGTAG CTGATTTTTC TGCAAGTGTT AAAAAGCCAC 1560  
 ATACCTTATT ACAGAGAGTC AAAGCCTGCA CAACAGAAGA GGATCAGGAG AAGCTGATGC 1620  
 AAATTACCAG TCTGCATTCA CTGAATGCCT TCTTGCTGCC AATTAAAACT GTAGGTGTGC 1680  
 AGGGTGACTG TCGTTCCTAC AGTTACGTGT GTGGAATCTC CAGTAAAGAT GAACCTGACT 1740  
 10 GGGAACTACT TATTTTCTG GCTAGGCTTA TACCTCGCAT GTGTCACAAC GTTAACAGAG 1800  
 TTGTTTATAT ATTTGGCCCA CCAGTTAAAG AACCTCCTAC AGATGTTACT CCCACTTTCT 1860  
 TGACAACAGG GGTGCTCAGT ACTTTACGCC AAGCTGATTT TGAGGCCCAT AACATTCTCA 1920  
 GGGAGTCTGG GTATGCTGGG AAAATCAGCC AGATGCCGGT GATTTTGACA CCATTACATT 1980  
 TTGATCGGGA CCCACTTCAA AAGCAGCCTT CATGCCAGAG ATCTGTGGTT ATTCGAACCT 2040  
 15 TTATTACTAG TGACTTCATG ACTGGTATAC CTGCAACACC TGGCAATGAG ATCCCTGTAG 2100  
 AGGTGGTATT AAAGATGGTC ACTGAGATTA AGAAGATTCC TGGTATTTCT CGAATTATGT 2160  
 ATGACTTAAC ATCAAAGCCC CCAGGAATA CTGAGTGGGA GTAATAAACT TC

Seq ID NO: 365 Protein sequence  
 Protein Accession #: AAA60331

1 11 21 31 41 51  
 | | | | |  
 MALCNGDSKL ENAGGDLKDG HHHYEGAVVI LDAGAQQYKV IDRRVRELTV QSEIFPLETP 60  
 25 AFAIKEQGFR AIIISGGPNS VYAEADAPWF PAIFTIGKPV LGICYGMQMM NKVFGGTVHK 120  
 KSVREDGVFN ISVDNTCSLF RGLQKEEVVL LTHGDSVDKV ADGFKVVAR S GNIVAGIANE 180  
 SKKLYGAQFH PEVGLTENGL VILKNFLYDI AGCSGTFVQ NRELECIREI KERVGTSKVL 240  
 VLLSGGVDST VCTALLNRAL NQEQVIAVHI DNGFMRKRES QSVEEALKKL GIQVKVINAA 300  
 HSFYNGTTTL PISDEDRTPR KRISKTLNMT TSPEEKRLII GDTFVKIANE VIGEMNLKPE 360  
 30 EVFLAQGLR PDLIESASLV ASGKAELIKT HHNDTELIRK LREEGKVIEP LKDFHKDEVR 420  
 ILGRELGLPE ELVSRHPFPG PGLAIRVICA EEPYICKDFP ETNNILKIVA DFSASVKKPH 480  
 TLLQRVKACT TEEDQEKLMQ ITSLHSLNAF LLPIKTVGVQ GDCRSYSYVC GISSKDEPDW 540  
 ESLIFLARLI PRMCHNVNRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADFEAHNLR 600  
 ESGYAGKISQ MPVILTPLHF DRDPLQKQPS CQRSVVIRTF ITSDFMGTGP ATPGNEIPVE 660  
 35 VVLKMTVEIK KIPGISRIMY DLTSPKPGTT EWE

Seq ID NO: 366 DNA sequence  
 Nucleic Acid Accession #: NM\_004219  
 Coding sequence: 46-654

1 11 21 31 41 51  
 | | | | |  
 GCGGCCCTCAG ATGAATGCGG CTGTTAAGAC CTGCAATAAT CCAGAATGGC TACTCTGATC 60  
 45 TATGTTGATA AGGAAAATGG AGAACCAGGC ACCCGTGTGG TTGCTAAGGA TGGGCTGAAG 120  
 CTGGGGTCTG GACCTTCAAT CAAAGCCTTA GATGGGAGAT CTCAAGTTTC AACACCACGT 180  
 TTTGGCAAAA CGTTCGATGC CCCACCAGCC TTACCTAAAG CTACTAGAAA GGCTTTGGGA 240  
 ACTGTCAACA GAGCTACAGA AAAGTCTGTA AAGACCAAGG GACCCCTCAA ACAAAAACAG 300  
 CCAAGCTTTT CTGCCAAAAA GATGACTGAG AAGACTGTTA AAGCAAAAAG CTCTGTTTCT 360  
 50 GCCTCAGATG ATGCCTATCC AGAAATAGAA AAATCTTTC CCTTCAATCC TCTAGACTTT 420  
 GAGAGTTTTG ACCTGCTTGA AGAGCACCAG ATTGCGCACC TCCCTTGGAG TGGAGTGCCT 480  
 CTCATGATCC TTGACGAGGA GAGAGAGCTT GAAAAGCTGT TTCAGCTGGG CCCCCCTTCA 540  
 CCTGTGAAGA TGCCCTCTCC ACCATGGGAA TCCAATCTGT TGCAGTCTCC TTCAAGCATT 600  
 CTGTCGACCC TGGATGTTGA ATTGCCACCT GTTTGCTGTG ACATAGATAT TTAAATTCT 660  
 55 TAGTGCTTCA GAGTTTGTGT GTATTTGTAT TAATAAAGCA TTCTTCAACA GAAAAAATAA 720  
 AAAAAA

Seq ID NO: 367 Protein sequence  
 Protein Accession #: NP\_004210

1 11 21 31 41 51  
 | | | | |  
 MATLIYVDKE NGEPTGRVVA KDGLKLGSFP SIKALDGRSQ VSTPRFGKTF DAPPALPKAT 60  
 65 RKALGTVNRA TEKSVKTKGP LKQKQPSFSA KKMTEKTVKA KSSVPASDDA YPEIEKFFPF 120  
 NPLDFESFDL PEEHQIAHLP LSGVPLMILD EERELEKLFQ LGPPSPVKMP SPPWESNLLQ 180  
 SPSSILSTLD VELPPVCCDI DI

Seq ID NO: 368 DNA sequence  
 Nucleic Acid Accession #: NM\_000597  
 Coding sequence: 118-1104

1 11 21 31 41 51  
 | | | | |  
 75 ATTCGGGGCG AGGGAGGAGG AAGAAGCGGA GGAGGCGGCT CCCGCTCGCA GGGCCGTGCA 60  
 CCTGCCCGCC CGCCGCTCG CTCGCTCGCC CGCCGCGCCG CGCTGCCGAC CGCCAGCATG 120  
 CTGCCGAGAG TGGGCTGCC CGCGCTGCC CTGCCGCCG CGCCGCTGCT GCCGCTGCTG 180  
 CCGCTGCTGC TGTGCTACT GGGCGCGAGT GGGCGCGGCG CGGGGGCGCG CGCGGAGGTG 240  
 CTGTTCCGCT GCGCGCCGTG CACACCCGAG CGCTTGGCCG CCTGCGGGCC CCGCGCGGTT 300  
 80 GCGCCGCCCC CGCGGTGCG CGCAGTGGCC GGAGGCGCCC GCATGCCATG CGCGGAGCTC 360  
 GTCCGGGAGC CGGGCTGCGG CTGCTGCTCG GTGTGCGCCC GGCTGGAGGG CGAGGCGTGC 420  
 GCGCTCTACA CCGCGCGCTG CGGCCAGGGG CTGCGCTGCT ATCCCAACCC GGGCTCCGAG 480  
 CTGCCCCCTG AGGCGCTGGT CATGGGCGAG GGCATTGTG AGAAGCGCCG GGACGCCGAG 540  
 TATGGCGCCA GCGCGGAGCA GGTTCAGAG AATGGCGATG ACCACTCAGA AGGAGGCCTG 600  
 85 ATGGAGAACC ACGTGGACAG CACCATGAAC ATGTTGGGCG GGGGAGGCG TGCTGGCCGG 660  
 AAGCCCTCA AGTCGGGTAT GAAGGAGCTG CCGGTGTTCC GGGAGAAGGT CACTGAGCAG 720  
 CACCGGCAGA TGGGCAAGGG TGGCAAGCAT CACCTTGGCC TGGAGGAGCC CAAGAAGCTG 780  
 CGACCACCCC CTGCCAGGAC TCCCTGCCAA CAGGAAGTGG ACCAGGTCCT GGAGCGGATC 840



TCCACCATGC GCCTTCCGGA TGAGCGGGGC CCTCTGGAGC ACCTCTACTC CCTGCACATC 900  
 CCCAACTGTG ACAAGCATGG CCTGTACAAC CTCAAACAGT GCAAGATGTC TCTGAACGGG 960  
 CAGCGTGGGG AGTGCTGGTG TGTGAACCCC AACACCGGGA AGCTGATCCA GGGAGCCCCC 1020  
 ACCATCCGGG GGGACCCCGA GTGTCACTCT TTCTACAATG AGCAGCAGGA GGCTTGCGGG 1080  
 GTGCACACCC AGCGGATGCA GTAGACCGCA GCCAGCCGGT GCCTGGCGCC CCTGCCCCC 1140  
 GCCCCTCTCC AAACACCGGC AGAAAACGGA GAGTGCTTGG GTGGTGGGTG CTGGAGGATT 1200  
 TTCCAGTTCT GACACACGTA TTTATATTG GAAAGAGACC AGCACCGAGC TCGGCACCTC 1260  
 CCCGGCCTCT CTCTTCCCAG CTGCAGATGC CACACCTGCT CTTTCTTGCT TTCCCCGGGG 1320  
 GAGGAAGGGG GTTGTGGTGC GGGAGCTGGG GTACAGGTTT GGGGAGGGGG AAGAGAAATT 1380  
 TTTATTTTGG AACCCCTGTG TCCCTTTTGC ATAAGATTAA AGGAAGGAAA AGT

Seq ID NO: 369 Protein sequence  
 Protein Accession #: NP\_000588

1 11 21 31 41 51  
 | | | | |  
 MLPRVGC PAL PLPPPPLLP LLLLLLLLGA SGGGGGARAE VLFRCPPCTP ERLAACGPPP 60  
 VAPPAVA AV AGGARMPCAE LVREPGCGCC SVCARLEGEA CGVYTPRCGQ GLRCYPHPGS 120  
 ELPLQALVMG EGTCEKRRDA EYGASPEQVA DNGDDHSEGG LVENHVDSTM NMLGGGGSAG 180  
 RKPLKSGMKE LAVFREKVT E QHRQMGKGGK HHLGLEEPKK LRPPPARTPC QQELDQVLER 240  
 ISTMRLPDER GPLEHLYSLH IPNCDKHGLY NLKQCKMSLN GQRGECWCVN PNTGKLIQGA 300  
 PTIRGDPECH LFYNEQQEAC GVHTQRMQ

Seq ID NO: 370 DNA sequence  
 Nucleic Acid Accession #: NM\_004264  
 Coding sequence: 6-440

1 11 21 31 41 51  
 | | | | |  
 GGAACATGGC GGATCGGCTC ACGCAGCTTC AGGACGCTGT GAATTCGCTT GCAGATCAGT 60  
 TTTGTAATGC CATTGGAGTA TTGCAGCAAT GTGCTCCTCC TGCTCTTTC AATAATATTC 120  
 AGACAGCAAT TAACAAAGAC CAGCCAGCTA ACCCTACAGA AGAGTATGCC CAGCTTTTTC 180  
 CAGCACTGAT TGCACGAACA GCAAAAGACA TTGATGTTTT GATAGATTCC TTACCCAGTG 240  
 AAGAATCTAC AGCTGCTTTA CAGGCTGCTA GCTTGTATAA GCTAGAAGAA GAAAACCATG 300  
 AAGCTGCTAC ATGTGTGGAG GATGTTGTTT ATCGAGGAGA CATGCTTCTG GAGAAGATAC 360  
 AAAGCGCACT TGCTGATATT GCACAGTCAC AGCTGAAGAC AAGAAGTGGT ACCCATAGCC 420  
 AGTCTCTTCC AGACTCATAG CATCAGTGGA TACCATGTGG CTGAGAAAAG AACTGTTTGA 480  
 GTGCCATTAA GAATCTGCA TCAGACTTAG ATACAAGCCT TACCAACAAT TACAGAAACA 540  
 TTAAACACTA TGACACATTA CCTTTTTCAG TATTTTAAAT AGTCTTCTAT TTCACTCTT 600  
 GATAAGCTTA TAAATCATGA TTGAATCAGC TTTAAAGCAT CATACCATCA TTTTAACT 660  
 GAGTGAAATT ATTAAGGCAT GTAATACATT AATGAACATA ATATAAGGAA ACATATGTAA 720  
 AATTCTGTTA TGACATAATT TATGTCTCCA TTTTGTGTGA TTGGCCAGTA CTTTACAAAT 780  
 C

Seq ID NO: 371 Protein sequence  
 Protein Accession #: NP\_004255

1 11 21 31 41 51  
 | | | | |  
 MADRLTQLQD AVNSLADQFC NAIGVLQCCG PPASFNNIQT AINKDQPANP TEEYAQLFAA 60  
 LIARTAKDID VLIDSLPSEE STAALQAASL YKLEENHEA ATCVEDVVYR GDMLEKIQS 120  
 ALADIAQSQL KTRSGTHSQS LPDS

Seq ID NO: 372 DNA sequence  
 Nucleic Acid Accession #: AJ271091  
 Coding sequence: 1-1113

1 11 21 31 41 51  
 | | | | |  
 ATGGAGAATC AGGTGTTGAC GCCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60  
 CTGCGCGTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120  
 CATTTCAAAG CTCAAGGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180  
 TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAACTGA CCCAGAGGCA GGTAAACATT 240  
 ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300  
 CTGTTTTTGG CTCCTGACTT TGATCGTTGG CTGGATGAAT CTGATGCGGA AATGGAGCTC 360  
 AGAGCTAAGG AAGAAGAGCG CCTAAATAAA CTCCGACTGG AAAGCGAAGG CTCTCCTGAA 420  
 ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGTGCA ATTCTTGGGA 480  
 TTCTCCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCCTTT 540  
 TATGACACAT TCCATACTGT GGCCTGACATG ATGTATTTCT GCCAGATGCT GGCAGTTGTG 600  
 GAAACTATCA ATGCAGCAAT TGGAGTCACT ACGTCACCGG TGCTGCCTTC TCTGATCCAG 660  
 CTTCTTGGAA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGCAGAAC 720  
 AAAGCTGTGG TTTTCTTTGT GTTTTATTTG TGGAGTGCAA TTGAAATTTT CAGGTACTCT 780  
 TTCTACATGC TGACGTGCAT TGACATGGAT TGGGAAGGTGC TCACATGGCT TCGTTACACT 840  
 CTGTGGATTG CTTTATATCC ACTGGGATGT TTGGCGGAAG CTGTCTCAGT GATTCACTCC 900  
 ATTCCAATAT TCAATGAGAC CGGACGATTC GATTTTACAT TGCCATATCC AGTGAAAATC 960  
 AAAGTTAGAT TTTCTTTTAT TCTTCAGATT TATCTTATAT TGATATTTT AGGTTTATAT 1020  
 ATAAATTTTC GTTACCTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGGCT 1080  
 CATGCTGTG ATCCAGCGC TTTGGGAGGC TGA

Seq ID NO: 373 Protein sequence  
 Protein Accession #: CAB69070

1 11 21 31 41 51  
 | | | | |  
 MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60  
 FLDLVKPEPV YKLTQRQVNI TVQKKVSQWW ERLTKQEKRP LFLAPDFDRW LDESDAEMEL 120  
 RAKEERLNK LRLESEGSPE TLTNLRKGYL FMYNLVQFLG FSWIFVNLTV RFCILGKESF 180



YDTFHTVADM MYFCQMLAVV ETINAAIGVT TSPVLPSLIQ LLGRNFILFI IFGTMEEMQN 240  
KAVVFFVFYF WSAIEIFRYS FYMLTCIDMD WKVLTWLRYS LWIPLYPLGC LAEAVSVIQS 300  
IPIFNETGRF SFTLPYPVKI KVRFSFFLQI YLIMIFLGLY INFRHLYKQR RLKMRAGAVA 360  
HACDPSALGG

Seq ID NO: 374 DNA sequence  
Nucleic Acid Accession #: NM\_016395  
Coding sequence: 1-1113

1 11 21 31 41 51  
| | | | | |  
ATGGAGAATC AGGTGTTGAC GCCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60  
CTGCGCGTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120  
CATTTCAAAG CTCAAGGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180  
TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAAACTGA CCCAGAGGCA GGTAAACATT 240  
ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300  
CTGTTTTTGG CTCCTGACTT TGATCGTTGG CTGGATGAAT CTGATGCGGA AATGGAGCTC 360  
AGAGCTAAGG AAGAAGAGCG CCTAAATAAA CTCCGACTGG AAAGCGAAGG CTCTCCTGAA 420  
ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGTGCA ATTCTTGGGA 480  
TTCTCCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCCTTT 540  
TATGACACAT TCCATACTGT GGCTGACATG ATGTATTTCT GCCAGATGCT GGCAGTTGTG 600  
GAAACTATCA ATGCAGCAAT TGGAGTCACT ACCTCACCGG TGCTGCCTTC TCTGATCCAG 660  
CTTCTTGGA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGCAGAAC 720  
AAAGCTGTGG TTTTCTTTGT GTTTTATTTG TGGAGTGCAA TTGAAATTTT CAGGTACTCT 780  
TTCTACATGC TGACGTGCAT TGACATGGAT TGGAAAGGTG TCACATGGCT TCGTTACACT 840  
CTGTGGATTG CCTTATATCC ACTGGGATGT TTGGCGGAAG CTGTCTCAGT GATTCACTCC 900  
ATTCCAATAT TCAATGAGAC CGGACGATTC AGTTTCACAT TGCCATATCC AGTGAAAATC 960  
AAAGTTAGAT TTTCTTTTTT TCTTCAGATT TATCTTATAA TGATATTTTT AGGTTTATAC 1020  
ATAAATTTTC GTCACCTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGGCT 1080  
CATGCCTGTG ATCCCAGCGC TTTGGGAGGC TGA

Seq ID NO: 375 Protein sequence  
Protein Accession #: NP\_057479

1 11 21 31 41 51  
| | | | | |  
MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60  
FLDLVKPEPV YKLTQRQVNI TVQKKVSQWW ERLTKQEKRP LFLAPDFDRW LDESDAEMEL 120  
RAKEEERLNK LRLESEGSPE LRLNLRKGYL FMYNLVQFLG FSWIFVNLTV RFCILGKESF 180  
YDTFHTVADM MYFCQMLAVV ETINAAIGVT TSPVLPSLIQ LLGRNFILFI IFGTMEEMQN 240  
KAVVFFVFYF WSAIEIFRYS FYMLTCIDMD WKVLTWLRYS LWIPLYPLGC LVEAVSVIQS 300  
IPIFNETGRF SFTLPYPVKI KVRFSFFLQI YLIMIFLGLY INFRHLYKQR RRRYGGKKRKR 360  
STKKKDLDFG LPV

Seq ID NO: 376 DNA sequence  
Nucleic Acid Accession #: NM\_005987  
Coding sequence: 1-270

1 11 21 31 41 51  
| | | | | |  
ATGAATTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA GCAGCAGCAG 60  
GTGAAACAAC CTTGCCAGCC TCCACCCAG GAACCATGCA TCCCCAAAC CAAGGAGCCC 120  
TGCCAACCCA AGGTGCCTGA GCCCTGCCAC CCCAAAGTGC CTGAGCCCTG CCAGCCCAAG 180  
ATTCCAGAGC CCTGCCAGCC CAAGGTGCCT GAGCCCTGCC CTTCAACGGT CACTCCAGCA 240  
CCAGCCCAAG AGAAGACCAA GCAGAAGTAA

Seq ID NO: 377 Protein sequence  
Protein Accession #: NP\_005978

1 11 21 31 41 51  
| | | | | |  
MNSQQQKQPC TPPPQPPQQQ VKQPCQPPPO EPCIPKTKEP CQPKVPEPCH PKVPEPCQPK 60  
IPEPCQPKVP EPCPSTVTPA PAQQKTKQK

Seq ID NO: 378 DNA sequence  
Nucleic Acid Accession #: NM\_002105  
Coding sequence: 74-505

1 11 21 31 41 51  
| | | | | |  
ACAGCAGTTA CACTGCGGCG GCGTCTGTT CTAGTGTGTT AGCCGTCGTG CTTACCCGGT 60  
CTACCTCGCT AGCATGTCGG GCCGCGGCAA GACTGGCGGC AAGGCCCGCG CCAAGGCCAA 120  
GTGCGGCTCG TCGCGCGCCG GCCTCCAGTT CCCAGTGGGC CGTGTACACC GGCTGCTGCG 180  
GAAGGGCCAC TACGCGGAGC GCGTTGGCGT CGGCGCGCCA GTGTACCTGG CGGCAGTGCT 240  
GGAGTACCTC ACCGCTGAGA TCCTGGAGCT GCGGGGCAAT GCGGCCCGCG ACAACAAGAA 300  
GACGCGAATC ATCCCCCGCC ACCTGCAGCT GGCCATCCGC AACGACGAGG AGCTCAACAA 360  
GCTGCTGGGC GCGGTGACGA TCGCCAGGG AGGCGTCTG CCCAACATCC AGGCCGTGCT 420  
GCTGCCCAAG AAGACCAGCG CCACCGTGGG GCCGAAGGCG CCTCGGGCG GCAAGAAGGC 480  
CACCCAGGCC TCCCAGGAGT ACTAAGAGGG CCCGCGCCGC GGCCGGCCGC CCCAGCTCCC 540  
CATGCCACCA CAAAGGCCCT TTAAAGGGCC ACCACCGCCC TCATGGAAAG AGCTGAGCCG 600  
CTTCAGACTG CGGGGCAAGC GGGCCGCGG TCCCTTCCCC TCCCTCCCC TCGCCCGCCT 660  
TCGCGGCCCG GCCTCGAGTC GCGCCCGGCC CCCGCTCCCG TCCCGCACCG CCTGCCGCGT 720  
CGGCCTCGGG CCTGCCCTGT CCGCCGTCCG CCCTCCGGTA GGGTTCGGGC CTTCCGGATG 780  
CGGCTTGGGC GCTCTTCGGG GACCTCCGTG GCGCGGAAGA CCCGAGCCTG CCGGGGGGAG 840



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65

GCCGGCGGCG CCGCACCTGC CCGCCTCGGC GTTCGTGACT CAGCCGCCCC ATCCCCGAGTC 900  
GCTAAGGGGGC TCGGGGGAGG CCGCAGCACC TTCTGGAAGA CTTGGCCTTC CGCTCTGACG 960  
CAGGGCCGAG GTGGGCAGTC CAGGCCGAGA GCCGGCGGCC CTGAAGGTGA GTGAGGCCCT 1020  
CGGCAGCTGC AGCCGGGGTG TCTGGTACCC CCCC GGCGTG GTGCTTAGCC CAGGACTTTC 1080  
AGACGGCCGC TGGCCGGGAG GCTTTGGTGG GAGAGACGCG ATCGCCGATT TCGGTCTGGC 1140  
GCCCCCTTCTG CCGCCGGGAC CCAGGCCTTT CACATCAGCT CTCCCTCCAT CTTTCATTCAT 1200  
AGGTCTGCGC TGGGGCCGGG ACGAAGCACT TGGTAACAGG CACATCTTCC TCCCGAGTGA 1260  
CTGCCTCCTA GGAGGACATT TAGGGGAGGG CAGAGGCCCTG CAGTTTGGCT TCACGGCTGG 1320  
CTATGTGGAC AGCAAGAGTC GTTTTGCAGA ACGCGACTGG CAGCCAGGCC TGTCGGGGCC 1380  
CCGACGCCGC CCCATTTCCT TTCCAGCAAA CTCAACTCGG CAATCCAAGC ACCTAGATAC 1440  
CAGCACAAGT CGGTTAATCC CTGTCTGGAC TGAGCCTCCG TTGGCTTCTG AACTGGAATT 1500  
CTGCAGCTAA CCCTTCCACG ACTAGAACCT TAGGCATTGG GGAGTTTATG ATGGACTAAT 1560  
TTTATTAAAG GATTGTTTTT TTTTT

Seq ID NO: 379 Protein sequence  
Protein Accession #: NP\_002096

1 11 21 31 41 51  
| | | | |  
MSGRGKTGGK ARAKAKSRSS RAGLQFPVGR VHRLLRKHY AERVGAGAPV YLAADVLEYLT 60  
AEILELAGNA ARDNKKTRII PRHLQLAIRN DEELNKLGG VTIAQGGVLP NIQAVLLPKK 120  
TSATVGPAP SGGKKATQAS QEY

Seq ID NO: 380 DNA sequence  
Nucleic Acid Accession #: AL136942  
Coding sequence: 184-864

1 11 21 31 41 51  
| | | | |  
ACGCGTCCGG CAGAAGCTCG GAGCTCTCGG GGTATCGAGG AGGCAGGCCC GCGGGCGCAC 60  
GGGCGAGCGG GCCGGGAGCC GGAGCGGCGG AGGAGCCGGC AGCAGCGGCG CGGCGGGCTC 120  
CAGGCGAGGC GGTGACGCT CCTGAAAACCT TCGCGCGCGC CTCGCGCCAC TCGCGCCGGA 180  
GCGATGAAGA TGGTCGCGCC CTGGACGCGG TTCTACTCCA ACAGCTGCTG CTTGTGCTGC 240  
CATGTCCGCA CCGGCACCAT CCTGCTCGGC GTCTGGTATC TGATCATCAA TGCTGTGGTA 300  
CTGTTGATTT TATTGAGTGC CCTGGCTGAT CCGGATCAGT ATAACCTTTC AAGTTCTGAA 360  
CTGGGAGGTG ACTTTGAGTT CATGGATGAT GCCAACATGT GCATTGCCAT TGCGATTTCT 420  
CTTCTCATGA TCCTGATATG TGCTATGGCT ACTTACGGAG CGTACAAGCA ACGCGCAGCC 480  
TGGATCATCC CATTCTTCTG TTACCAGATC TTTGACTTTG CCCTGAACAT GTTGGTTGCA 540  
ATCACTGTGC TTATTATCC AACTCCATT CAGGAATACA TACGGCAACT GCCTCCTAAT 600  
TTTCCCTACA GAGATGATGT CATGTCAGTG AATCCTACCT GTTGGTCCCT TATTATTCTT 660  
CTGTTTATTA GCATTATCTT GACTTTTAAG GGTACTTGA TTAGCTGTGT TTGGAAGTGC 720  
TACCGATACA TCAATGGTAG GAACTCCTCT GATGTCCTGG TTTATGTTAC CAGCAATGAC 780  
ACTACGGTGC TGCTACCCCC GTATGATGAT GCCACTGTGA ATGGTGCTGC CAAGGAGCCA 840  
CCGCCACCTT ACGTGTCTGC CTAAGCCTTC AAGTGGGCGG AGCTGAGGGC AGCAGCTTGA 900  
CTTTCAGAGC ATCTGAGCAA TAGTCTGTGT ATTTCACTTT TGCCATGAGC CTCTCTGAGC 960  
TTGTTTGTG CTGAAATGCT ACTTTTAAA ATTTAGATGT TAGATTGAAA ACTGTAGTTT 1020  
TCAACATATG CTTTGCTAGA ACACTGTGAT AGATTAACTG TAGAATTCTT CCTGTACGAT 1080  
TGGGGATATA ACGGGCTTCA CTAACCTTCC CTAGGCATTG AAACCTCCCC CAAATCTGAT 1140  
GGACCTAGAA GTCTGCTTTT GTACCTGCTG GGCCCCAAAG TTGGGCATTT TTCTCTCTGT 1200  
TCCCTCTCTT TTGAAAATGT AAAATAAAAC CAAAAATAGA CAACCTTTTC TTCAGCCATT 1260  
CCAGCATAGA GAACAAAACC TTATGGAAAC AGGAATGTCA ATTGTGTAAT CATTGTTCTA 1320  
ATTAGGTAAA TAGAAGTCTT TATGTATGTG TTACAAGAAT TTCCCCCACA ACATCCTTTA 1380  
TGACTGAAGT TCAATGACAG TTTGTGTTTG GTGGTAAAGG ATTTTCTCCA TGGCCTGAAT 1440  
TAAGACCATT AGAAAGCACC AGGCCGTGGG AGCAGTGACC ATCTACTGAC TGTCTTTGTG 1500  
GATCTTGTGT CCAGGGACAT GGGGTGACAT GCCTCGTATG TGTTAGAGGG TGGAATGGAT 1560  
GTGTTTGGCG CTGCATGGGA TCTGGTGCCC CTCTTCTCCT GGATTCACAT CCCCACCCAG 1620  
GGCCCGCTTT TACTAAGTGT TCTGCCCTAG ATTGGTTCAA GGAGGTCATC CAACTGACTT 1680  
TATCAAGTGG AATTGGGATA TATTGATAT ACTTCTGCCT AACAACATGG AAAAGGGTTT 1740  
TCTTTTCCCT GCAAGCTACA TCCTACTGCT TTGAACTTCC AAGTATGTCT AGTCACCTTT 1800  
TAAAATGTAA ACATTTTCAG AAAAATGAGG ATTGCCTTCC TTGTATGCGC TTTTACCTT 1860  
GACTACCTGA ATTGCAAGGG ATTTTATAT ATTCAATATG TACAAAGTCA GCAACTCTCC 1920  
TGTTGGTTCA TTATTGAATG TGCTGTAAT TAAGTCGTTT GCAATTAAAA CAAGGTTTGC 1980  
CCACATCCAA AAAAAAAAAA AAAAA

Seq ID NO: 381 Protein sequence  
Protein Accession #: CAB66876

1 11 21 31 41 51  
| | | | |  
MKMVAPWTRF YSNSCCLCCH VRTGTILLGV WYLIINAVVL LILLSALADP DQYNFSSSEL 60  
GGDFEFMDDA NMCIAIAISL LMILICAMAT YGAYKQRAAW IIPFFCYQIF DFALNMLVAI 120  
TVLIYPNSIQ EYIRQLPPNF PYRDDVMSVN PTCLVLIILL FISIILTFKG YLISCVWNCY 180  
RYINGRNSSD VLVYVTSNDT TVLLPPYDDA TVNGAAKEPP PPYVSA

Seq ID NO: 382 DNA sequence  
Nucleic Acid Accession #: NM\_002510  
Coding sequence: 92-1774

1 11 21 31 41 51  
| | | | |  
CAGATGCCAG AAGAACACTG TTGCTCTTGG TGGACGGGCC CAGAGGAATT CAGAGTTAAA 60  
CCTTGAGTGC CTGCGTCCGT GAGAATTGAT CATGAAATGT CTCTACTATT TCCTGGGATT 120  
TCTGCTCCTG GCTGCAAGAT TGCCACTTGA TGCCGCCAAA CGATTTTCATG ATGTGCTGGG 180  
CAATGAAAGA CTTTCTGCTT ACATGAGGGA GCACAATCAA TTAAATGGCT GGTCTTCTGA 240  
TGAAAATGAC TGGAATGAAA AACTCTACCC AGTGTGGAAG CGGGGAGACA TGAGGTGGAA 300  
AAACTCCTGG AAGGGAGGCC GTGTGCAGGC GGTCTTGACC AGTGACTCAC CAGCCCTCGT 360



5	GGGCTCAAAT	ATAACATTTG	CGGTGAACCT	GATATTCCCT	AGATGCCAAA	AGGAAGATGC	420
	CAATGGCAAC	ATAGTCTATG	AGAAGAATCG	CAGAAATGAG	GCTGGTTTAT	CTGCTGATCC	480
	ATATGTTTAC	AACTGGACAG	CATGGTCAGA	GGACAGTGAC	GGGGAAAAATG	GCACCGGCCA	540
	AAGCCATCAT	AACGTCTTCC	CTGATGGGAA	ACCTTTTCCCT	CACCACCCCG	GATGGAGAAG	600
	ATGGAATTTT	ATCTACGTCT	TCCACACACT	TGGTCAGTAT	TTCCAGAAAT	TGGGACGATG	660
10	TTTCAGTGAGA	GTTTCTGTGA	ACACAGCCAA	TGTGACACTT	GGGCCTCAAC	TCATGGAAGT	720
	GACTGTCTAC	AGAAGACATG	GACGGGCATA	TGTTCCCATC	GCACAAGTGA	AAGATGTGTA	780
	CGTGGTAACA	GATCAGATTG	CTGTGTTTGT	GACTATGTTT	CAGAAGAACG	ATCGAAATTC	840
	ATCCGACGAA	ACCTTCCTCA	AAGATCTCCC	CATTATGTTT	GATGTCCTGA	TTTCATGATCC	900
	TAGCCACTTC	CTCAATTATT	CTACCATTAA	CTACAAGTGG	AGCTTCGGGG	ATAATACTGG	960
15	CCTGTTTGTG	TCCACCAATC	ATACTGTGAA	TCACACGTAT	GTGCTCAATG	GAACCTTCAG	1020
	CCTTAACCTC	ACTGTGAAAG	CTGCAGCACC	AGGACCTTGT	CCGCCACCGC	CACCACCACC	1080
	CAGACCTTCA	AAACCCACCC	CTTCTTTAGG	ACCTGCTGGT	GACAACCCCC	TGGAGCTGAG	1140
	TAGGATTCCT	GATGAAAAC	GCCAGATTAA	CAGATATGGC	CACCTTCAAG	CCACCATCAC	1200
	AATTGTAGAG	GGTAATCTTAG	AGGTTAACAT	CATCCAGATG	ACAGACGTCC	TGATGCCGGT	1260
20	GCCATGGCCT	GAAAGCTCCC	TAATAGACTT	TGTCGTGACC	TGCCAAGGGA	GCATTCCAC	1320
	GGAGGTCTGT	ACCATCATTT	CTGACCCAC	CTGCGAGATC	ACCCAGAACA	CAGTCTGCAG	1380
	CCCTGTGGAT	GTGGATGAGA	TGTGTCTGCT	GACTGTGAGA	CGAACCTTCA	ATGGGTCTGG	1440
	GACGTACTGT	GTGAACCTCA	CCCTGGGGGA	TGACACAAGC	CTGGCTCTCA	CGAGCACCTT	1500
	GATTTCTGTT	CCTGACAGAG	ACCCAGCCTC	GCCTTTAAGG	ATGGCAAACA	GTGCCCTGAT	1560
25	CTCCGTTGGC	TGCTTGGCCA	TATTTGTCAC	TGTGATCTCC	CTCTTGGTGT	ACAAAAAACA	1620
	CAAGGAATAC	AACCAATAG	AAAATAGTCC	TGGGAATGTG	GTCAGAAGCA	AAGGCCTGAG	1680
	TGTCTTTCTC	AACCGTGCAA	AAGCCGTGTT	CTTCCCGGGA	AACCAGGAAA	AGGATCCGCT	1740
	ACTCAAAAAC	CAAGAATTTA	AAGGAGTTTC	TTAAATTTTCG	ACCTTGTTC	TGAAGCTCAC	1800
	TTTTTCAGTGC	CATTGATGTG	AGATGTGCTG	GAGTGGCTAT	TAACCTTTTT	TTCTTAAAGA	1860
30	TTATTGTTAA	ATAGATATTG	TGGTTTGGGG	AAGTTGAATT	TTTTATAGGT	TAAATGTCAT	1920
	TTTAGAGATG	GGGAGAGGGA	TTTACTGCA	GGCAGCTTCA	GCCATGTTGT	GAAACTGATA	1980
	AAAGCAACTT	AGCAAGGCTT	CTTTTCATTA	TTTTTTATGT	TTCACTTATA	AAGTCTTAGG	2040
	TAAGTAGTAG	GATAGAAAACA	CTGTGTCCCG	AGAGTAAGGA	GAGAAGCTAC	TATTGATTAG	2100
	AGCCTAACCC	AGGTAACTG	CAAGAAGAGG	CGGGATACTT	TCAGCTTCC	ATGTAAGTGT	2160
35	ATGCATAAAG	CCATGTAGT	CCAGTTTCTA	AGATCATGTT	CCAAGCTAAC	TGAATCCAC	2220
	TTCAATACAC	ACTCATGAAC	TCTGATGGA	ACAATAACAG	GCCCAAGCCT	GTGGTATGAT	2280
	GTGCACACTT	GCTAGACTCA	GAATAAATAC	TACTCTCATA	AATGGGTGGG	AGTATTTTGG	2340
	TGACAACCTA	CTTTGCTTGG	CTGAGTGAAG	GAATGATATT	CATATATTCA	TTTATTCCAT	2400
	GGACATTTAG	TTAGTGCTTT	TTATATACCA	GGCATGATGC	TGAGTGACAC	TCTTGTGTAT	2460
40	ATTTCCAAAT	TTTTGTATAG	TCGCTGCACA	TATTTGAAAT	CATATATTAA	GACTTTCCAA	2520
	AGATGAGGTC	CCTGGTTTTT	CATGGCAACT	TGATCAGTAA	GGATTTCAAC	TCTGTTTGTA	2580
	ACTAAAACCA	TCTACTATAT	GTTAGACATG	ACATTCTTTT	TCTCTCCTTC	CTGAAAAATA	2640
	AAGTGTGGGA	AGAGACAAAA	AAAAAAAAAA				

Seq ID NO: 383 Protein sequence  
Protein Accession #: NP\_002501

45	1	11	21	31	41	51	
	MECLYYFLGF	LLLAARLPLD	AAKRFHDLV	NERPSAYMRE	HNQLNGWSSD	ENDWNEKLYP	60
	VWKRGDNRWK	NSWKGRVQA	VLTSDSPALV	GSNITFAVNL	IFPRCQKEDA	NGNIVYEKNC	120
	RNEAGLSADP	YVYNWTAWSE	DSDGENGTGQ	SHHNVPDGGK	PFPHPGWRR	WNFIYVFHTL	180
50	GQYFQKLGR	SVRVSVNTAN	VTLGPQLMEV	TVYRRHGRAY	VPIAQVKDVG	VVTDQIPVFG	240
	TMFQKNDRNS	SDETFLKDL	IMFDVLIHDP	SHFLNYSTIN	YKWSFGDNTG	LFVSTNHTVN	300
	HTYVLNGTFS	LNLTVKAAAP	GPCPPPPPPP	RPSKPTPSLG	PAGDNPLELS	RIPDENCQIN	360
	RYGHFQATIT	IVEGILEVNI	IQMTDVLMPV	PWPESSLIDF	VVTCQGS IPT	EVCTIISDPT	420
	CEITQNTVCS	PVDVDEMCLL	TVRRFTFNGS	TYCVNLTLGD	DTSLALTSTL	ISVPDRDPAS	480
55	PLRMANSALI	SVGCLAIFVT	VISLLVYKKH	KEYNPIENSP	GNVVRSGGLS	VFLNRAKAVF	540
	FPGNQEKDPL	LKNQEFKGV					

Seq ID NO: 384 DNA sequence  
Nucleic Acid Accession #: NM\_001134  
Coding sequence: 48-1877

60	1	11	21	31	41	51	
	TCCATATTGT	GCTTCCACCA	CTGCCAATAA	CAAAATAACT	AGCAACCATG	AAGTGGGTGG	60
	AATCAATTTT	TTTAATTTTC	CTACTAAATC	TTACTGAATC	CAGAACACTG	CATAGAAATG	120
	AATATGGAAT	AGCTTCCATA	TTGGATTCTT	ACCAATGTAC	TGCAGAGATA	AGTTTAGCTG	180
65	ACCTGGCTAC	CATATTTTTC	GCCAGTTTTC	TTCAAGAAGC	CACTTACAAG	GAAGTAAGCA	240
	AAATGGTGAA	AGATGCATTG	ACTGCAATTG	AGAAACCCAC	TGGAGATGAA	CAGTCTTCAG	300
	GGTGTTTAGA	AAACCAGCTA	CCTGCCTTTC	TGGAAGAAGT	TTGCCATGAG	AAAGAAATTT	360
	TGGAGAAGTA	CGGACATTCA	GACTGCTGCA	GCCAAAGTGA	AGAGGGAAGA	CATAACTGTT	420
	TTCTTGACAA	CAAAAAGCCC	ACTCCAGCAT	CGATCCCACT	TTTCCAAGTT	CCAGAACCTG	480
70	TCACAAGCTG	TGAAGCATAT	GAAGAAGACA	GGGAGACATT	CATGAACAAA	TTTATTATATG	540
	AGATAGCAAG	AAGGCATCCC	TTCTGTATG	CACCTACAAT	TCTTCTTTGG	GCTGCTCGCT	600
	ATGACAAAAT	AATTCATCT	TGCTGCAAAAG	CTGAAAATGC	AGTTGAATGC	TTCCAAACAA	660
	AGGCAGCAAC	AGTTACAAAA	GAATTAAGAG	AAAGCAGCTT	GTTAAATCAA	CATGCATGTG	720
	CAGTAATGAA	AAATTTTGGG	ACCCGAAGCTT	TCCAAGCCAT	AACTGTTACT	AAACTGAGTC	780
75	AGAAGTTTAC	CAAAGTTAAT	TTACTGAAA	TCCAGAAACT	AGTCTTGGAT	GTGGCCCATG	840
	TACATGAGCA	CTGTTGCAGA	GGAGATGTGC	TGGATTGTCT	GCAGGATGGG	GAATAAATCA	900
	TGTCCTACAT	ATGTTCTCAA	CAAGACACTC	TGTCAAACAA	AATAACAGAA	TGCTGCAAAAC	960
	TGACCACGCT	GGAACGTGGT	CAATGTATAA	TTTCATGCAGA	AAATGATGAA	AAACCTGAAG	1020
	GTCTATCTCC	AAATCTAAAC	AGGTTTTTAG	GAGATAGAGA	TTTTAACCAA	TTTCTTTCAG	1080
80	GGGAAAAAAA	TATCTTCTTG	GCAAGTTTTG	TTTCATGAATA	TTCAAGAAGA	CATCCTCAGC	1140
	TTGCTGTCTC	AGTAATCTTA	AGAGTTGCTA	AAGGATACCA	GGAGTTATTG	GAGAAGTGTT	1200
	TCCAGACTGA	AAACCTCTTT	GAATGCCAAG	ATAAGAGGAG	AGAAGAATTA	CAGAAATACA	1260
	TCCAGGAGAG	CCAAGCATTG	GCAAAGCGAA	GCTGCGGCTT	CTTCCAGAAA	CTAGGAGAAAT	1320
	ATTACTTACA	AAATGCGTTT	CTCGTTGCTT	ACACAAAGAA	AGCCCCCAG	CTGACCTCGT	1380
85	CGGAGCTGAT	GGCCATCACC	AGAAAAATGG	CAGCCACAGC	AGCCACTTGT	TGCCAACTCA	1440
	GTGAGGACAA	ACTATTGGCC	TGTGGCGGAG	GAGCGGCTGA	CATTATTATC	GGACACTTAT	1500



GTATCAGACA TGAAATGACT CCAGTAAACC CTGGTGTGG CCAGTGCTGC ACTTCTTCAT 1560  
ATGCCAACAG GAGGCCATGC TTCAGCAGCT TGGTGGTGGG TGAAACATAT GTCCCTCCTG 1620  
CATCTCTGA TGACAAGTTC ATTTTCCATA AGGATCTGTG CCAAGCTCAG GGTGTAGCGC 1680  
TGCAAACGAT GAAGCAAGAG TTTCTCATT ACCTTGTGAA GCAAAAGCCA CAAATAACAG 1740  
AGGAACAAC TGAAGCTGTC ATTGCAGATT TCTCAGGCCT GTTGGAGAAA TGCTGCCAAG 1800  
GCCAGGAACA GGAAGTCTGC TTTGCTGAAG AGGACAAAA ACTGATTTCA AAAACTCGTG 1860  
CTGCTTTGGG AGTTTAAATT ACTTCAGGGG AAGAGAAGAC AAAACGAGTC TTTCAATCGG 1920  
TGTGAACCTT TCTCTTTAAT TTTAACTGAT TTAACACTTT TTGTGAATTA ATGAAATGAT 1980  
AAAGACTTTT ATGTGAGATT TCCTTATCAC AGAAATAAAA TATCTCCAAA TG

Seq ID NO: 385 Protein sequence  
Protein Accession #: NP\_001125

1 11 21 31 41 51  
| | | | |  
MKWVESIFLI FLLNFESRT LHRNEYGIAS ILDSYQCTAE ISLADLATIF FAQFVQEATY 60  
KEVSKMVKDA LTAIEKPTGD EQSSGCLNQ LPAFLLEELCH EKEILEKYGH SDCCSQSEEG 120  
RHNCFLAHKK PTPASIPLFQ VPEPVTSCFA YEEDRETFMN KFIYEIARRH PFLYAPTILL 180  
WAARYDKIIP SCCKAENAVE CFQTKAATVT KELRESSLLN QHACAVMKNF GTRTFQAITV 240  
TKLSQKFTKV NFTEIQKLVL DVAHVHEHCC RGDVLDCLQD GEKIMSYICS QQDTLSNKIT 300  
ECCKLTTLER QCIIHAEND EKPEGLSPNL NRFLGDRDFN QFSSGEKNIF LASFVHEYSR 360  
RHPQLAVSVI LRVAKGYQEL LEKCFQTENP LECQDKGEEE LQKYIQESQA LAKRSCGLFQ 420  
KLGEYYLQNA FLVAYTKKAP QLTSSSELMAL TRKMAATAAT CCQLSEDKLL ACCEGAADII 480  
IGHLCIRHEM TPVNPVGVC CTSSYANRRP CFSSLVVDDE YVPPAFSDDK FIFHKDLCOA 540  
QGVALLQTMKQ EFLINLVKQK PQITEEQLEA VIADFSGLLE KCCQGEQEV CFAEEGQKLI 600  
SKTRALGV

Seq ID NO: 386 DNA sequence  
Nucleic Acid Accession #: NM\_002205.1  
Coding sequence: 1..3149

1 11 21 31 41 51  
| | | | |  
ATGGGGAGCC GGACGCCAGA GTCCCTCTC CACGCCGTGC AGCTGCGCTG GGGCCCCCGG 60  
CGCCGACCCC CGCTSSTGCC GCTGCTGTTG CTGCTSSTGC CGCCGCCACC CAGGGTCGGG 120  
GGCTTCAACT TAGACGCGGA GCGCCAGCA GTACTCTCGG GCGCCCGGG CTCTTCTTC 180  
GGATTCTCAG TGGAGTTTAA CCGGCCGGGA ACAGACGGGG TCAGTGTGCT GGTGGGAGCA 240  
CCCAAGGCTA ATACCAGCCA GCCAGGAGTG CTGCAGGGTG GTGCTGTCTA CCTCTGTCT 300  
TGGGGTGCCA GCGCCACACA GTGCACCCCA ATTGAATTG ACAGCAAAGG CTCTCGGCTC 360  
CTGGAGTCCCT CACTGTCCAG CTCAGAGGGA GAGGAGCCTG TGGAGTACAA GTCCTTGCA 420  
TGTTCCGGGG CAACAGTTTC AGCCCATGGC TCCTCCATCT TGGCATGCGC TCCACTGTAC 480  
AGCTGGCGCA CAGAGAAGGA GCCACTGAGC GACCCCGTGG GCACCTGCTA CCTCTCCACA 540  
GATAACTTCA CCCGAATTCT GGAGTATGCA CCCTGCCGCT CAGATTTTCAG CTGGGCAGCA 600  
GGACAGGGTT ACTGCCAAGG AGGCTTCAGT GCCGAGTTCA CCAAGACTGG CCGTGTGGTT 660  
TTAGGTGGAC CAGGAAGCTA TTTCTGGCAA GGCCAGATCC TGTCTGCCAC TCAGGAGCAG 720  
ATTGCAGAAT CTTATTACCC CGATTACCTG ATCAACCTGG TTCAGGGGCA GCTGCAGACT 780  
CGCCAGGCCA GTTCCATCTA TGATGACAGC TACCTAGGAT ACTCTGTGGC TGTGGGTGAA 840  
TTCAGTGGTG ATGACACAGA AGACTTTGTT GCTGGTGTGC CCAAGGGGAA CCTCACTTAC 900  
GGCTATGTCA CCATCCTTAA TGGCTCAGAC ATTCGATCCC TCTACAACCT CTCAGGGGAA 960  
CAGATGGCCT CCTACTTTGG CTATGCAGTG GCGGCCACAG ACGTCAATGG GGACGGGCTG 1020  
GATGACTTGC TGGTGGGGGC ACCCTGTCTC ATGGATCGGA CCCCTGACGG GCGGCCTCAG 1080  
GAGGTGGGCA GGGTCTACGT CTACCTGACG CACCCAGCCG GCATAGAGCC CACGCCACC 1140  
CTTACCCTCA CTGGCCATGA TGAGTTTGGC CGATTGGCA GCTCCTTGAC CCCCTGGGG 1200  
GACCTGGACC AGGATGGCTA CAATGATGTG GCCATCGGGG CTCCCTTTGG TGGGGAGACC 1260  
CAGCAGGGAG TAGTGTGTTG ATTTCTGGG GGCCAGGAG GGCTGGGCTC TAAGCCTTCC 1320  
CAGGTCTGTC AGCCCTGTG GGCAGCCAGC CACACCCAG ACTTCTTTGG CTCTGCCCTT 1380  
CGAGGAGGCC GAGACCTGGA TGGCAATGGA TATCCTGATC TGATTGTGGG GTCCTTTGGT 1440  
GTGGACAAGG CTGTGGTATA CAGGGGCGCG CCCATCGTGT CCGCTAGTGC CTCCCTCACC 1500  
ATCTTCCCGG CCATGTTCAA CCCAGAGGAG CCGAGCTGCA GCTTAGAGGG GAACCCGTG 1560  
GCCTGCATCA ACCTTAGCTT CTGCCTCAAT GCTTCTGGAA AACACGTTGC TGACTCCATT 1620  
GGTTTCACAG TGGAACTTCA GCTGGACTGG CAGAAGCAGA AGGGAGGGGT ACGGCGGGCA 1680  
CTGTTCTTGG CCTCCAGGCA GGCAACCCTG ACCCAGACCC TGCTCATCCA GAATGGGGCT 1740  
CGAGAGGATT GCAGAGAGAT GAAGATCTAC CTCAGGAACG AGTCAGAATT TCGAGACAAA 1800  
CTCTCGCCGA TTCACATCGC TCTCAACTTC TCCTTGACC CCCAAGCCCC AGTGGACAGC 1860  
CACGGCCTCA GGCCAGCCCT ACATTATCAG AGCAAGAGCC GGATAGAGGA CAAGGCTCAG 1920  
ATCTTGCTGG ACTGTGGAGA AGACAACATC TGTGTGCCCT ACCTGCAGCT GGAAGTGT 1980  
GGGGAGCAGA ACCATGTGTA CCTGGGTGAC AAGAATGCCC TGAACCTCAC TTTCCATGCC 2040  
CAGAATGTGG GTGAGGGTGG CGCCTATGAG GCTGAGCTTC GGGTCACCGC CCCTCCAGAG 2100  
GCTGAGTACT CAGGACTCGT CAGACACCCA GGGAACCTCT CCAGCCTGAG CTGTGACTAC 2160  
TTTGCCGTGA ACCAGAGCCG CCTGCTGGTG TGTGACCTGG GCAACCCCAT GAAGGCAGGA 2220  
GCCAGTCTGT GGGGTGGCCT TCGGTTTACA GTCCCTCATC TCCGGGACAC TAAGAAAACC 2280  
ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTCGCAAAG CGACGTGGTT 2340  
TCCTTTCGGC TCTCCGTGGA GGCTCAGGCC CAGGTCACCC TGAACGGTGT CTCCAAGCCT 2400  
GAGGCAGTGC TATTTCCAGT AAGCGACTGG CATCCCCGAG ACCAGCCTCA GAAGGAGGAG 2460  
GACCTGGGAC CTGCTGTCCA CCATGTCTAT GAGCTCATCA ACCAAGGCCC CAGCTCCATT 2520  
AGCCAGGGTG TGCTGGAAC TCACTGTCCC CAGGCTCTGG AAGGTGAGCA GCTCTTATAT 2580  
GTGACCAGAG TTACGGGACT CAACTGCACC ACCAATCACC CCATTAACCC AAAGGGCCTG 2640  
GAGTTGGATC CCGAGGGTTC CCTGCACCAC CAGCAAAAAC GGGAGAGCTC AAGCCGAGC 2700  
TCTGCTTCTT CCGGACCTCA GATCCTGAAA TGCCCGGAGG CTGAGTGT 2760  
TGTGAGCTCG GGCCCTGCA CCAACAAGAG AGCCAAAGTC TGCAGTTGCA TTTCCAGT 2820  
TGGGCCAAGA CTTTCTTGCA GCGGGAGCAC CAGCCATTTA GCCTGCAGTG TGAGGCTGT 2880  
TACAAAGCCC TGAAGATGCC CTACCGAATC CTGCCTCGGC AGCTGCCCCA AAAAGAGCGT 2940  
CAGGTGGCCA CAGCTGTGCA ATGGACCAAG GCAGAAGGCA GCTATGGCGT CCCACTGTGG 3000  
ATCATCATCC TAGCCATCCT GTTGGCCTC CTGCTCCTAG GTCTACTCAT CTACATCCTC 3060  
TACAAGCTTG GATTCTTCAA ACGCTCCCTC CCATATGGCA CCGCCATGGA AAAAGCTCAG 3120  
CTCAAGCCTC CAGCCACCTC TGATGCCTGA



Seq ID NO: 387 Protein sequence  
Protein Accession #: NP\_002196.1

5 1 11 21 31 41 51  
| | | | | |  
MGSRTPEspl HAVQLRWGPR RRPPLLPLLL LLLPPPPRVG GFNLDAEAPA VLSGPPGSFF 60  
GFSVEFYRPG TDGVSVLVGA PKANTSQPGV LQGGAVYLCp WGASPTQCTP IEFDSKGSRL 120  
LESSLSSSEG EEPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPVGTCTYLS 180  
10 DNFTRILEYA PCRSDFSWAA GQGYCQGGFS AEFTKTGRVV LGGPGSYFWQ GQILSATQEQ 240  
IAESYYPEYL INLVQGQLQT RQASSIYDDS YLGYSVAVGE FSGDDTEDFV AGVPKGNLTY 300  
GYVTILNGSD IRSLYNFSGE QMASYFGYAV AATDVNGDGL DDLLVGAPLL MDRTPDGRPQ 360  
EVGRVYVYLQ HPAGIEPTPT LTLTGHDEFG RFGSSLTPLG DLDQDGYNDV AIGAPFGGET 420  
QGGVVFVFPg GPGGLGSKPS QVLQPLWAAS HTPDFFGSAL RGGRLDLGNG YPDLIVGSFG 480  
15 VDKAVVYRGR PIVSASASLT IFPAMFNPEE RSCSLEGNPV ACINLSFCLN ASGKHVADSI 540  
GFTVELQLDW QKQKGGVRRR LFLASRQATL TQTLLIQNGA REDCREMKIY LRNESEFRDK 600  
LSPIHIALNF SLDPQAPVDS HGLRPALHYQ SKSRIEDKAQ ILLDCGEDNI CVPDLQLEVF 660  
GEQNHVYLGd KNALNLTfHA QNVGEGGAYE AELRVtAPPE AEYSGLVrHP GNFSSLSCDY 720  
FAVNQSRLLV CDLGNPMKAG ASLWGGLRFT VPHLRDTKKT IQFDFQILSK NLNNSQSDVv 780  
20 SFRLSVEAQA QVTlNGVSKP EAVLFPVSDW HPRDQPQKEE DLGPAVHHVY ELINQGPSSI 840  
SQGVLELSCP CPEAEQQLLY VTRVTGLNCT TNHPINPKGL ELDPEGSLHH QQKREAPSRs 900  
SASSGPIILK CPEAEQQLLY VTRVTGLNCT TNHPINPKGL ELDPEGSLHH QQKREAPSRs 960  
YKALKMPYRI LPRQLPQKER QVATAVQWTK AEGSYGVPLW IILAILFGL LLLGLLIYIL 1020  
YKLGFFKRSL PYGTAMEKAQ LKPPATSDA

Seq ID NO: 388 DNA sequence  
Nucleic Acid Accession #: NM\_002425  
Coding sequence: 26..1453

30 1 11 21 31 41 51  
| | | | | |  
AAAGAAGGTA AGGGCAGTGA GAATGATGCA TCTTGcATTC CTTGTGCTGT TGTGTCTGCC 60  
AGTCTGCTCT GCCTATCCTC TGAGTGGGGG AGCAAAAGAG GAGGACTCCA ACAAGGATCT 120  
TGCCCGACAA TACCTAGAAA AGTACTACAA CCTCGAAAAG GATGTGAAAC AGTTTAGAAG 180  
35 AAAGGACAGT AATCTCATTG TTAATAAAAT CCAAGGAATG CAGAAGTTCC TTGGGTTGGA 240  
GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGGAGT 300  
TCCTGACGTT GGTCACTTCA GCTCCTTTCC TGGCATGCCG AAGTGGAGGA AAACCCACCT 360  
TACATACAGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTTG ATTCTGCCAT 420  
TGAGAAAGCT CTGAAAGTCT GGGAAAGAGT GACTCCACTC ACATTCTCCA GGCTGTATGA 480  
40 AGGAGAGGCT GATATAATGA TCTCTTTGCG AGTTAAAGAA CATGGAGACT TTTACTCTTT 540  
TGATGGCCCA GGACACAGTT TGGCTCATGC CTACCCACCT GGACCTGGGC TTTATGGAGA 600  
TATTCACCTT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATTCCT 660  
CGTTGCTGCT CATGAACCTG GCCACTCCCT GGGGCTCTTT CACTCAGCCA AACTGAAGC 720  
TTTGATGTAC CCACTCTACA ACTCATTAC AGAGCTCGCC CAGTTCGCC TTTGCAAGA 780  
45 TGATGTGAAT GGCACTCAGT CTCTCTACGG ACCTCCCCCT GCCTCTACTG AGGAACCCCT 840  
GGTGCCCA CAATCTGTTC CTTGGGATG TGAGATGCCA GCCAAGTGTG ATCCTGCTTT 900  
GTCCTTCGAT GCTATCAGCA CTCTGAGGG AGAATATCTG TTCTTTAAAG ACAGATATTT 960  
TTGGCGAAGA TCCCACTGGA ACCCTGAACC TGAATTTTCA TTGATTTCTG CATTTTGGCC 1020  
CTCTCTCCA TCATATTGG ATGCTGCATA TGAAGTTAAC AGCAGGGACA CCGTTTAT 1080  
50 TTTTAAAGGA AATGAGTTCT GGGCCATCAG AGGAAATGAG GTACAAGCAG GTTATCCAAG 1140  
AGGCATCCAT ACCCTGGGTT TTCCTCCAAC CATAAGGAAA ATTGATGCAG CTGTTTCTGA 1200  
CAAGGAAAAG AAGAAAACAT ACTTCTTTG ACCTGGAGAT TGTATGAAAA 1260  
TAGCCAGTCC ATGGAGCAAG GCTTCCCTAG AGCTAATAGCT GATGACTTTC CAGGAGTTGA 1320  
GCCTAAGGTT GATGCTGTAT TACAGGCATT TGGATTTTTC TACTTCTTCA GTGGATCATC 1380  
55 ACAGTTTGAG TTTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAGA GTAACAGCTG 1440  
GTTACATTGC TAGGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTTAAT AAATCTAATA 1500  
ATTATTCATC TAATGTATTA TGAGCCAAAA TGGTTAATTT TTCTGTCATG TTCTGTGACT 1560  
GAAGAAGATG AGCCTTGcAG ATATCTGCAT GTGTCATGAA GAATGTTTCT GGAATTCTTC 1620  
ACTTGCTTTT GAATGCACT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680  
60 ATGTATTTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGGCCTGTTC 1740  
CTT

Seq ID NO: 389 Protein sequence  
Protein Accession #: NP\_002416

65 1 11 21 31 41 51  
| | | | | |  
MHLAFLVLLC LPVCSAYPLS GAAKEEDSNK DLAQQYLEKY YNLEKDVkQF RRKDSNLIVK 60  
KIQMQKFLG LEVTGKLDTD TLEVMRKPRC GVPDVGHFSS FPGMPKWRKT HLTyRIVNYT 120  
70 PDLPRDAVDS AIEKALKVWE EVTPLTFsRL YEGEADIMIS FAVKEHGDFY SFDGPGHSLA 180  
HAYPPGPGLY GDIHFDDDEK WTEDASGTNL FLVAAHELGH SLGLFHSANT EALMYPLYNs 240  
FTELAQFRLS QDDVNGIQSL YGPPPAStEE PLVPTKSVPS GSEMPAKCDP ALSFDAISTL 300  
RGEYLFFKDR YFWRRSHWNP EPEFHLISAF WPSLPSYLDa AYEVRNSRDtV FIFKGNEFWA 360  
75 IRGNEVQAGY PRGIHTLGFP PTIRKIDAAV SDKEKKKTYF FAADKYWRFd ENSQSMEQGF 420  
PRLIADDFPG VEPKVDAVLQ AFGFFYFFSG SSQFEFDpNA RMVTHILKSN SWLHC

Seq ID NO: 390 DNA sequence  
Nucleic Acid Accession #: NM\_002421.2  
Coding sequence: 1..1409

80 1 11 21 31 41 51  
| | | | | |  
ATGCACAGCT TTCCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTc ACACAGCTTC 60  
CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120  
85 TACTACAACC TGAAGAATGA TGGGAGGCAA GTTGAAAAGC GGAGAAATAG TGGCCCACTG 180  
GTTGAAAAAT TGAAGCAAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCAGAT 240  
GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCTGATGT GGCTCAGTTT 300



5  
10  
15  
20

```
GTCCTCACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAAT 360
TACACGCCAG ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTCCCAACTC 420
TGGAGTAATG TCACACCTCT GACATTCACC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480
ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540
CTTGCTCATG CTTTTCAACC AGGCCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600
GAAAGGTGGA CCAACAATTT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGAACTC 660
GGCCATTCTC TTGGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCCTAGCTAC 720
ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATTG ATGGCATCCA AGCCATATAT 780
GGACGTTCCC AAAATCCTGT CCAGCCCATC GGCCCAAAA CCCCCAAAAGC ATGTGACAGT 840
AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900
TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTTCAT TTCTGTTTTT 960
TGGCCACAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020
CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140
CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT 1200
GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCCT 1260
GGAATTGGCC ACAAAGTTGA TGCAGTTTTT ATGAAAGATG GATTTTTTCTA TTTCTTTCAT 1320
GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG
```

Seq ID NO: 391 Protein sequence  
Protein Accession #: NP\_002412.1

25  
30  
35

```
1 11 21 31 41 51
| | | | |
MHSFPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLKQMGEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYRIEN 120
YTPDLPRADV DHAIEKAFQL WSNVTPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180
LAHAFQPGPG IGGDAHFDED ERWTNNFREY NLHRVAAHEL GHSLGLSHST DIGALMYPSY 240
TFSGDVQLAQ DDIDIGIAIY GRSQNPVQPI GPQTPKACDS KLTFDAITTI RGEVMFFKDR 300
FYMRTNPFYP EVELNFISVF WPQLPNGLEA AYEFAADRDEV RFFKGNKYWA VQGQNVLHGY 360
PKDIYSSFGF PRTVKHIDAA LSEENTGKTY FFVANKYWRY DEYKRSMDDPG YPKMIAHDFP 420
GIGHKVDAVF MKDGGFFFFH GTRQYKFDPK TKRILTQKA NSWFNCRKN
```

Seq ID NO: 392 DNA sequence  
Nucleic Acid Accession #: NM\_002421.2  
Coding sequence: 1..1409

40  
45  
50  
55  
60  
65

```
1 11 21 31 41 51
| | | | |
ATGCACAGCT TTCCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTC ACACAGCTTC 60
CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120
TACTACAACC TGAAGAATGA TGGGAGGCAA GTTGAAAAGC GGAGAAATAG TGGCCCACTG 180
GTTGAAAAAT TGAAGCAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCAGAT 240
GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCTGATGT GGCTCAGTTT 300
GTCCTCACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAAT 360
TACACGCCAG ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTCCCAACTC 420
TGGAGTAATG TCACACCTCT GACATTCACC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480
ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540
CTTGCTCATG CTTTTCAACC AGGCCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600
GAAAGGTGGA CCAACAATTT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGCCCTC 660
GGCCATTCTC TTGGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCCTAGCTAC 720
ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATTG ATGGCATCCA AGCCATATAT 780
GGACGTTCCC AAAATCCTGT CCAGCCCATC GGCCCAAAA CCCCCAAAAGC ATGTGACAGT 840
AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900
TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTTCAT TTCTGTTTTT 960
TGGCCACAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020
CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140
CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT 1200
GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCCT 1260
GGAATTGGCC ACAAAGTTGA TGCAGTTTTT ATGAAAGATG GATTTTTTCTA TTTCTTTCAT 1320
GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG
```

Seq ID NO: 393 Protein sequence  
Protein Accession #: NP\_002412.1

70  
75  
80

```
1 11 21 31 41 51
| | | | |
MHSFPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLKQMGEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYRIEN 120
YTPDLPRADV DHAIEKAFQL WSNVTPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180
LAHAFQPGPG IGGDAHFDED ERWTNNFREY NLHRVAAHAL GHSLGLSHST DIGALMYPSY 240
TFSGDVQLAQ DDIDIGIAIY GRSQNPVQPI GPQTPKACDS KLTFDAITTI RGEVMFFKDR 300
FYMRTNPFYP EVELNFISVF WPQLPNGLEA AYEFAADRDEV RFFKGNKYWA VQGQNVLHGY 360
PKDIYSSFGF PRTVKHIDAA LSEENTGKTY FFVANKYWRY DEYKRSMDDPG YPKMIAHDFP 420
GIGHKVDAVF MKDGGFFFFH GTRQYKFDPK TKRILTQKA NSWFNCRKN
```

Seq ID NO: 394 DNA sequence  
Nucleic Acid Accession #: NM\_014331.2  
Coding sequence: 1..1506

85

```
1 11 21 31 41 51
| | | | |
```



	ATGGTCAGAA	AGCCTGTTGT	GTCCACCATC	TCCAAAGGAG	GTTACCTGCA	GGGAAATGTT	60
	AACGGGAGGC	TGCCTTCCCT	GGGCAACAAG	GAGCCACCTG	GGCAGGAGAA	AGTGCAGCTG	120
	AAGAGGAAAG	TCACCTTACT	GAGGGGAGTC	TCCATTATCA	TTGGCACCAT	CATTGGAGCA	180
5	GGAAATCTCA	TCTCTCTAA	GGGCGTGCTC	CAGAACACGG	GCAGCGTGGG	CATGTCTCTG	240
	ACCATCTGGA	CGGTGTGTGG	GGTCCTGTCA	CTATTTGGAG	CTTTGTCTTA	TGCTGAATTG	300
	GGAAACAATA	TAAAGAAATC	TGGAGGTCAT	TACACATATA	TTTTGGAAAGT	CTTTGGTCCA	360
	TTACCAGCTT	TTGTACGAGT	CTGGGTGGAA	CTCCTCATAA	TACGCCCTGC	AGCTACTGCT	420
	GTGATATCCC	TGGCATTGCG	ACGCTACATT	CTGGAACCAT	TTTTTATTCA	ATGTGAAATC	480
10	CCTGAACCTG	CGATCAAGCT	CATTACAGCT	GTGGGCATAA	CTGTAGTGAT	GGTCCTAAAT	540
	AGCATGAGTG	TCAGCTGGAG	CGCCCGGATC	CAGATTTTCT	TAACCTTTTG	CAAGCTCACA	600
	GCAATTCTGA	TAATTATAGT	CCCTGGAGTT	ATGCAGCTAA	TTAAAGGTCA	AACGCAGAAC	660
	TTTAAAGACG	CGTTTTTCAGG	AAGAGATTCA	AGTATTACGC	GGTTGCCACT	GGCTTTTTTAT	720
	TATGGAATGT	ATGCATATGC	TGGCTGGTTT	TACCTCAACT	TTGTTACTGA	AGAAGTAGAA	780
	AACCCTGAAA	AAACCATTCC	CCTTGCAATA	TGTATATCCA	TGGCCATTGT	CACCATTGGC	840
15	TATGTGCTGA	CAAATGTGGC	CTACTTTACG	ACCATTAAATG	CTGAGGAGCT	GCTGCTTTCA	900
	AATGCAGTGG	CAGTGACCTT	TTCTGAGCGG	CTACTGGGAA	ATTTCTCATT	AGCAGTTCCG	960
	ATCTTTGTTG	CCCTCTCCTG	CTTTGGCTCC	ATGAACGGTG	GTGTGTTTGC	TGTCTCCAGG	1020
	TTATTCTATG	TTGCGTCTCG	AGAGGGTCAC	CTTCCAGAAA	TCCTCTCCAT	GATTCATGTC	1080
20	CGCAAGCACA	CTCCTCTACC	AGCTGTTATT	GTTTTGCACC	CTTTGACAAT	GATAATGCTC	1140
	TTCTCTGGAG	ACCTCGACAG	TCTTTTGAAT	TTCTCAGTT	TTGCCAGGTG	GCTTTTTTATT	1200
	GGGCTGGCAG	TTGCTGGGCT	GATTTATCTT	CGATACAAAT	GCCCAGATAT	GCATCGTCTC	1260
	TTCAAGGTGC	CACGTGTCAT	CCCAGCTTTG	TTTTCTTCA	CATGCCTCTT	CATGGTTGCC	1320
	CTTTCCCTCT	ATTCGGACCC	ATTTAGTACA	GGGATTGGCT	TCGTCTATC	TCTGACTGGA	1380
25	GTCCCTGCGT	ATTATCTCTT	TATTATATGG	GACAAGAAAC	CCAGGTGGTT	TAGAATAATG	1440
	TCAGAGAAAA	TAACCAGAAC	ATTACAAATA	ATACTGGAAG	TTGTACCAGA	AGAAGATAAG	1500
	TTATGAACTA	ATGGACTTGA	GATCTTGGCA	ATCTGCCCAA	GGGGAGACAC	AAAATAGGGA	1560
	TTTTTACTTC	ATTTTCTGAA	AGTCTAGAGA	ATTACAACCT	TGGTGATAAA	CAAAGGAGT	1620
	CAGTTAATTT	TATTCATATA	TTTTAGCAT	TTTGAACATA	TTTCTAAGAA	ATTTAGTTAT	1680
30	AACTCTATGT	AGTTATAGAA	AGTGAATATG	CAGTTATTCT	ATGAGTCGCA	CAATTCTTGA	1740
	GTCTCTGATA	CCTACCTATT	GGGGTTAGGA	GAAAAGACTA	GACAATTACT	ATGTGGTCAT	1800
	TCTCTACAA	ATATGTTAGC	ACGGCAAAGA	ACCTTCAAAT	TGAAGACTGA	GATTTTTCTG	1860
	TATATATGGG	TTTTGTAAAG	ATGGTTTTAC	ACACTACAGA	TGTCTATACT	GTGAAAAGTG	1920
	TTTTCAATTC	TGAAAAAAG	CATACATCAT	GATTATGGCA	AAGAGGAGAG	AAAGAAATTT	1980
35	ATTTTACATT	GACATTGCAT	TGCTTCCCT	TAGATACCAA	TTTAGATAAC	AAACACTCAT	2040
	GCTTTAATGG	ATTATACCCA	GAGCACTTTG	AACAAAGGTC	AGTGGGGATT	GTGAATACA	2100
	TTAAAGAAGA	GTTTCTAGGG	GCTACTGTTT	ATGAGACACA	TCCAGGAGTT	ATGTTTAAGT	2160
	AAAAATCCTT	GAGAATTTAT	TATGTCAGAT	GTTTTTTCAT	TCATTATCAG	GAAGTTTTAG	2220
	TTATCTGTCA	TTTTTTTTTT	TCACATCAGT	TTGATCAGGA	AAGTGTATAA	CACATCTTAG	2280
40	AGCAAGAGTT	AGTTTGGTAT	TAAATCCTCA	TTAGAACAAC	CACCTGTTTC	ACTAATAACT	2340
	TACCCTTGAT	GAGTCTATCT	TAACATATGC	ATTTTAAGCC	TTCAAATTAC	ATTATCAACA	2400
	TGAGAGAAAT	AACCAACAAA	GAAGATGTTT	AAAATAATAG	TCCCATATCT	GTAATCATAT	2460
	CTACATGCAA	TGTTAGTAAT	TCTGAAGTTT	TTTAAATTTA	TGGCTATTTT	TACACGATGA	2520
	TGAATTTTGA	CAGTTTGTGC	ATTTTCTTTA	TACATTTTAT	ATTCTTCTGT	TAAAATATCT	2580
45	CTTCAGATGA	AACTGTCCAG	ATTAATTAGG	AAAAGGCATA	TATTAACATA	AAAATTGCAA	2640
	AAGAAATGTC	GCTGTAAATA	AGATTTACAA	CTGATGTTTC	TAGAAAATTT	CCACTTCTAT	2700
	ATCTAGGCTT	TGTCAGTAAT	TTCCACACCT	TAATTATCAT	TCAACTTGCA	AAAGAGACAA	2760
	CTGATAAGAA	GAAAATTGAA	ATGAGACACT	TGGGATAAGT	GTTTGTGTTT	AGAAGATGTT	2820
	GTTTTCGCCAG	TATTAGAAAA	TACTGTGAGC	CGGGCATGGT	GGCTTACATC	TGTAATCCCA	2880
50	GCACTTTGGG	AGGCTGAGGG	GGTGGATCAC	CTGAGGTCGG	GAGTTCTAGA	CCAGCCTGAC	2940
	CAACATGGAG	AAACCCCATC	TCTACTAAAA	ATACAAAATT	AGCTGGGCAT	GGTGGCACAT	3000
	GCTGGTAATC	TCAGCTATTG	AGGAGGCTGA	GGCAGGAGAA	TTGCTTGAAC	CCGGGAGGCG	3060
	GAGGTTGCAG	TGAGCCAAGA	TTGCACCCT	GTACTCCAGC	CTGGGTGACA	AAGTCAGACT	3120
	CCATCTCCAA	AAAAAAAAAA	AAAA				

Seq ID NO: 395 Protein sequence  
Protein Accession #: NP\_055146.1

	1	11	21	31	41	51	
60	MVRKPVVSTI	SKGGYLQGNV	NGRLPSLGNK	EPFGQEKVQL	KRKVTLLRGV	SIIIGTIIGA	60
	GIFISPKGVL	QNTGSVGMSL	TIWTVCGVLS	LFGALSYAEL	GTTIKKSGGH	YTYILEVFGP	120
	LPAFVRVWVE	LLIIRPAATA	VISLAFGRYI	LEPFFIQCEI	PELAIKLITA	VGITVVMVLN	180
	SMSVSWASRI	QIFLTFCKLT	AILIIVPGV	MQLIKGQTQN	FKDAFSGRDS	SITRLPLAFY	240
65	YGMAYAGWF	YLNFTVEEVE	NPEKTIPLAI	CISMAITIGV	YVLTNVAYFT	TINAEELLS	300
	NAVAVTFSEF	LLGNFSLAVP	IFGLNFCFGS	MNGGVFAVSR	LFYVASREGH	LPEILSMIHV	360
	RKHTPLPAVI	VLHPLTMIML	FSGLDLSLLN	FLSFARWLFI	GLAVAGLIYL	RYKCPDMHRP	420
	FKVPLFIPAL	FSFTCLFMVA	LSLYSDPFST	GIGFVITLTG	VPAYYLFIIW	DKKPRWFRIM	480
	SEKITRTLQI	ILEVVPEEDK	L				

Seq ID NO: 396 DNA sequence  
Nucleic Acid Accession #: NM\_006528  
Coding sequence: 57..764

	1	11	21	31	41	51	
75	GCCGCCAGCG	GCTTTCTCGG	ACGCCTTGCC	CAGCGGGCCG	CCCGACCCCC	TGCACCATGG	60
	ACCCCGCTCG	CCCCCTGGGG	CTGTGCTGTT	TGCTGCTTTT	CCTGACGGAG	GCTGCACTGG	120
80	GCGATGCTGC	TCAGGAGCCA	ACAGGAAATA	ACGCGGAGAT	CTGTCTCCTG	CCCCTAGACT	180
	ACGGACCCTG	CCGGGCCCTA	CTTCTCCGTT	ACTACTACGA	CAGGTACACG	CAGAGCTGCC	240
	GCCAGTTCCT	GTACGGGGGC	TGCGAGGGCA	ACGCCAACAA	TTTCTACACC	TGGGAGGCTT	300
	GCGACGATGC	TTGCTGGAGG	ATAGAAAAAG	TTCCCAAAGT	TTGCCGGCTG	CAAGTGAGTG	360
	TGGACGACCA	GTGTGAGGGG	TCCACAGAAA	AGTATTCTTT	TAATCTAAGT	TCCATGACAT	420
85	GTGAAAAATT	CTTTTCCGGT	GGGATGCACC	GGAAACGGAT	TGAGAACAGG	TTTCCAGATG	480
	AAGCTACTTG	TATGGGCTTC	TGCGCACCAA	AGAAAAATCC	ATCATTTTGC	TACAGTCCAA	540
	AAGATGAGGG	ACTGTGCTCT	GCCAATGTGA	CTCGCTATTA	TTTTAATCCA	AGATACAGAA	600
	CCTGTGATGC	TTTCACCTAT	ACTGGCTGTG	GAGGGAATGA	CAATAACTTT	GTTAGCAGGG	660



AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGAAAAAGAA AAAGAAGATG CCAAAGCTTC 720  
GCTTTGCCAG TAGAATCCCG AAAATTCGGA AGAAGCAATT TTAAACATTC TTAATATGTC 780  
ATCTTGTGTTG TCTTTATGGC TTATTTGCCT TTATGGTTGT ATCTGAAGAA TAATATGACA 840  
GCATGAGGAA ACAAATCATT GGTGATTTAT TCACCAGTTT TTATTAATAC AAGTCACTTT 900  
TTCAAAAATT TGGATTTTAT TATATATAAC TAGCTGCTAT TCAAAATGTGA GTCTACCATT 960  
TTTAATTTAT GGTTCAACTG TTTGTGAGAG GAATTCCTGC AATGCATAAG ATATAAAAGC 1020  
AAATATGACT CACTCATTTT TGGGGTCTG ATTCTTGATT TCAGAAGAGG ATCATAACTG 1080  
AAACAACATA AGACAATATA ATCATGTGCT TTTAACATAT TTGAGAATAA AAAGGACTAG 1140  
CC

Seq ID NO: 397 Protein sequence  
Protein Accession #: NP\_006519

1 11 21 31 41 51  
| | | | |  
MDPARPLGLS ILLFLTEAA LGDAAQEPTG NNAEICLLPL DYGPCRALLL RYYYDRYTQS 60  
CRQFLYGGCE GNANFYWE ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFNLSSM 120  
TCEKFFSGGC HRNRIENRFP DEATCMGFCA PKKIPSFCYS PKDEGLCSAN VTRYFNPRI 180  
RTCDAFTYTG CGGNDNNFVS REDCKRACAK ALKKKKKMPK LRFASIRIKI RKKQF

Seq ID NO: 398 DNA sequence  
Nucleic Acid Accession #: NM\_001508.1  
Coding sequence: 1..1361

1 11 21 31 41 51  
| | | | |  
ATGGCTTCAC CCAGCCTCCC GGGCAGTGAC TGCTCCCAAA TCATTGATCA CAGTCATGTC 60  
CCCAGATTTG AGGTGGCCAC CTGGATCAAA ATCACCCTTA TTCTGGTGTA CCTGATCATC 120  
TTCGTGATGG GCCTTCTGGG GAACAGCGTC ACCATTCCGG TCACCCAGGT GCTGCAGAAG 180  
AAAGGATACT TGCAGAAGGA GGTGACAGAC CACATGGTGA GTTTGGCTTG CTCGGACATC 240  
TTGGTGTTC TCATCGGCAT GCCCATGGAG TTCTACAGCA TCATCTGGAA TCCCCTGACC 300  
ACGTCCAGCT ACACCTGTCT CTGCAAGCTG CACACTTTCC TCTTCGAGGC CTGCAGCTAC 360  
GCTACGCTGC TGCACGTGCT GACGCTCAGC TTTGAGCGCT ACATCGCCAT CTGTACCCCC 420  
TTCAGGTACA AGGCTGTGTC GGGACCTTGC CAGGTGAAGC TGCTGATTGG CTTCTCTGG 480  
GTCACCTCCG CCCTGGTGGC ACTGCCCTTG CTGTTTGCCA TGGGTACTGA GTACCCCTG 540  
GTGAACGTGC CCAGCCACCG GGGTCTCACT TGCAACCGCT CCAGCACCCG CCACCACGAG 600  
CAGCCCGAGA CCTCCAATAT GTCCATCTGT ACCAACCTCT CCAGCCGCTG GACCGTGTTC 660  
CAGTCCAGCA TCTTCGGCGC CTTCTGTGCT TACCTCGTGG TCCTGCTCTC CGTAGCCTTC 720  
ATGTGCTGGA ACATGATGCA GGTGCTCATG AAAAGCCAGA AGGGCTCGCT GGCCGGGGGC 780  
ACGCGGCCTC CGCAGCTGAG GAAGTCCGAG AGCGAAGAGA GCAGGACCGC CAGGAGGCAG 840  
ACCATCATCT TCCTGAGGCT GATTGTGTG ACATTGGCCG TATGCTGGAT GCCCAACCAG 900  
ATTCGGAGGA TCATGGCTGC GGCCAAACCC AAGCAGACT GGACGAGGTC CTAATTCCGG 960  
GCGTACATGA TCCTCCTCCC CTTCTCGGAG ACGTTTTTCT ACCTCAGCTC GGTCAACAAC 1020  
CCGCTCCTGT ACACGGTGTCT CTCGACGAGC TTTCCGGCGG TGTTCTGTCA GGTGCTGTGC 1080  
TGCCGCTGT CGCTGCAGCA CGCCAAACCC GAGAAGCGCC TGCGCGTACA TGCGCACTCC 1140  
ACCACCGACA GCGCCCGCTT TGTGACGCGC CCGTTGCTCT TCGCGTCCCG GCGCCAGTCC 1200  
TCTGCAAGGA GAACTGAGAA GATTTTCTTA AGCACTTTTC AGAGCGAGGC CGAGCCCCAG 1260  
TCTAAGTCCC AGTCATTGAG TCTCGAGTCA CTAGAGCCCA ACTCAGGCGC GAAACCAGCC 1320  
AATTCTGCTG CAGAGAATGG TTTTCAGGAG CATGAAGTTT GA

Seq ID NO: 399 Protein sequence  
Protein Accession #: NP\_001499.1

1 11 21 31 41 51  
| | | | |  
MASPSLPGLS CSQIIDHSHV PEFEVATWIK ITLILVYLII FVMGLLGNSV TIRVTQVLQK 60  
KGYLQKEVTD HMYSLACSDI LVFLIGMPME FYSLIWNPLT TSSYTLSCKL HTFLFEACSY 120  
ATLLHVLTLF FERYIAICHP FRYKAVSGPC QVKLLIGFVW VTSALVALPL LFAMGTEYPL 180  
VNVPSHRGLT CNRSSTRHHE QPETSNSMIC TNLSSRWTFV QSSIFGAFV YLVVLLSVAF 240  
MCWNMMQVLM KSQKSLAGG TRPPQLRKSE SEESRTARRQ TIIFLRILIV TLAVCWMPNQ 300  
IRRIMAAAKP KHDWTRSYFR AYMILLPFSE TFFYLSSVIN PLLYTVSSQQ FRRVFVQVLC 360  
CRLSLQHANH EKRLRVHAHS TTDSARFVQR PLLFASRRQS SARRTEKIFL STFQSEAEFQ 420  
SKSQSLSLLES LEPNSGAKPA NSAAENGFOE HEV

Seq ID NO: 400 DNA sequence  
Nucleic Acid Accession #: NM\_006475.1  
Coding sequence: 28..2538

1 11 21 31 41 51  
| | | | |  
AACAGAACTG CAACGGAGAG ACTCAAGATG ATTCCCTTTT TACCCATGTT TTCTCTACTA 60  
TTGCTGCTTA TTGTTAACCC TATAAACGCC AACAATCATT ATGACAAGAT CTTGGCTCAT 120  
AGTCGTATCA GGGGTCGGGA CCAAGGCCCA AATGTCTGTG CCCTTCAACA GATTTTGGGC 180  
ACCAAAAAGA AATACTTCAG CACTTGTAAG AACTGGTATA AAAAGTCCAT CTGTGGACAG 240  
AAAACGACTG TTTTATATGA ATGTTGCCCT GGTATATGA GAATGGAAGG AATGAAAGGC 300  
TGCCGAGCAG TTTTGCCCAT TGACCATGTT TATGGCACTC TGGGCATCGT GGGAGCCACC 360  
ACAACGCAGC GCTATTCTGA CGCCTCAAAA CTGAGGGAGG AGATCGAGGG AAAGGGATCC 420  
TTCACCTACT TTGCACCGAG TAATGAGGCT TGGGACAACT TGGATTCTGA TATCCGTAGA 480  
GGTTTGAGAG GCAACGTGAA TGTTGAATTA CTGAATGCTT TACATAGTCA CATGATTAAT 540  
AAGAGAATGT TGACCAAGGA CTTAAAAAAT GGCATGATTA TTCCTTCAAT GTATAACAAT 600  
TTGGGGCTTT TCATTAACCA TTATCCTAAT GGGGTTGTCA CTGTTAATTG TGCTCGAATC 660  
ATCCATGGGA CCTAGATTGC AACAAATGGT GTTGTCATG TCATTGACCG TGTGCTTACA 720  
CAAATTGGTA CCTCAATTCA AGACTTCATT GAAGCAGAAAG ATGACCTTTC ATCTTTTAGA 780  
GCAGCTGCCA TCACATCGGA CATATTGGAG GCCCTTGGA GAGACGGTCA CTTCACTC 840  
TTTGCTCCCA CCAATGAGGC TTTTGAGAAA CTTCCACGAG GTGTCCTAGA AAGGTTTCATG 900  
GGAGACAAAG TGGCTTCCGA AGCTCTTATG AAGTACCACA TCTTAAATAC TCTCCAGTGT 960  
TCTGAGTCTA TTATGGGAGG AGCAGTCTTT GAGACGCTGG AAGGAAATAC AATTGAGATA 1020



5  
10  
15  
20  
25  
30  
35  
40

```

GGATGTGACG GTGACAGTAT AACAGTAAAT GGAATCAAAA TGGTGAACAA AAAGGATATT 1080
GTGACAAATA ATGGTGTGAT CCATTTGATT GATCAGGTCC TAATTCCTGA TTCTGCCAAA 1140
CAAGTTATTG AGCTGGCTGG AAAACAGCAA ACCACCTTCA CGGATCTTGT GGCCCAATTA 1200
GGCTTGGCAT CTGCTCTGAG GCCAGATGGA GAATACACTT TGCTGGCACC TGTGAATAAT 1260
GCATTTTCTG ATGATACTCT CAGCATGGTT CAGCGCCTCC TTAAATTAAT TCTGCAGAAT 1320
CACATATTGA AAGTAAAAGT TGGCCTTAAT GAGCTTTACA ACGGGCAAAT ACTGGAAACC 1380
ATCGGAGGCA AACAGCTCAG AGTCTTCGTA TATCGTACAG CTGTCTGCAT TGAAAATTCA 1440
TGCATGGAGA AAGGGAGTAA GCAAGGGAGA AACGGTGCAG TTCACATATT CCGCGAGATC 1500
ATCAAGCCAG CAGAGAAATC CCTCCATGAA AAGTTAAAAC AAGATAAGCG CTTTAGCACC 1560
TTCCTCAGCC TACTTGAAGC TGCAGACTTG AAAGAGCTCC TGACACAACC TGGAGACTGG 1620
ACATTATTTG TGCCAACCAA TGATGCTTTT AAGGGAATGA CTAGTGAAGA AAAAGAAATT 1680
CTGATACGGG ACAAAAATGC TCTTCAAAAC ATCATTCTTT ATCACCTGAC ACCAGGAGTT 1740
TTCATTGGAA AAGGATTTGA ACCTGGTGTT ACTAACATTT TAAAGACCAC ACAAGGAAGC 1800
AAAACTTTT TGAAAGAAGT AAATGATACA CTTCTGGTGA ATGAATTGAA ATCAAAAGAA 1860
TCTGACATCA TGACAACAAA TGGTGTAAAT CATGTTGTAG ATAAACTCCT CTATCCAGCA 1920
GACACACCTG TTGGAATGA TCAACTGCTG GAAATACTTA ATAAATTAAT CAAATACATC 1980
CAAAATTAAGT TTGTTCTGTT TAGCACCTTC AAAGAAATCC CCGTGAAGTGT CTATACAACT 2040
AAAATTATAA CCAAAGTTGT GGAACCAAAA ATTAAAGTGA TTGAAGGCAG TCTTCAGCCT 2100
ATTATCAAAA CTGAAGGACC CACACTAACA AAAGTCAAAA TTGAAGGTGA ACCTGAATTC 2160
AGACTGATTA AAGAAGGTGA AACAATAACT GAAGTGATCC ATGGAGAGCC AATTATTAAA 2220
AAATACACCA AAATCATTGA TGGAGTGCCT GTGGAATAAA CTGAAAAAGA GACACGAGAA 2280
GAACGAATCA TTACAGGTCC TGAAATAAAA TACACTAGGA TTTCTACTGG AGGTGGAGAA 2340
ACAGAAGAAA CTCTGAAGAA ATTGTTACAA GAAGAGGTCA CCAAGGTCAC CAAATTCATT 2400
GAAGGTGGTG ATGGTCATTT ATTTGAAGAT GAAGAAATTA AAAGACTGCT TCAGGGAGAC 2460
ACACCCGTGA GGAAGTTGCA AGCCAACAAA AAAGTTCAAG GTTCTAGAAG ACGATTAAGG 2520
GAAGGTCGTT CTCAGTGAAA ATCCAAAAAC CAGAAAAAAA TGTTTATACA ACCCTAAGTC 2580
AATAACCTGA CCTTAGAAAA TTGTGAGAGC TCGGAACTG TCAGGAAGT AAACATCAGC 2640
ACAAAGAAGC AATCATCAA TAATCTGAA CACAAATTTA ATATTTTTTT TTCTGAATGA 2700
GAAACATGAG GGAATTTGTG GAGTTAGCCT CCTGTGGTAA AGGAATTGAA GAAAATATAA 2760
CACCTTACAC CCTTTTTTCAT CTTGACATTA AAAGTTCTGG CTAACCTTGG AATCCATTAG 2820
AGAAAAATCC TTGTCACCAG ATTCATTACA ATTCAAATCG AAGAGTTGTG AACTGTTATC 2880
CCATTGAAAA GACCGAGCCT TGTATGTATG TTATGGATAC ATAAATGCA CGCAAGCCAT 2940
TATCTCTCCA TGGGAAGCTA AGTTATAAAA ATAGGTGCTT GGTGTACAAA ACTTTTATA 3000
TCAAAAGGCT TTGCACATTT CTATATGAGT GGGTTTACTG GTAAATTATG TTATTTTTTA 3060
CAACTAATTT TGTACTCTCA GAATGTTTGT CATATGCTTC TTGCAATGCA TATTTTTTAA 3120
TCTCAAACGT TTCAATAAAA CCATTTTTCA GATATAAAGA GAATTACTTC AAATTGAGTA 3180
ATTCAGAAAA ACTCAAGATT TAAGTTAAAA AGTGGTTTGG ACTTGGGAA

```

Seq ID NO: 401 Protein sequence  
Protein Accession #: NP\_006466.1

45  
50  
55  
60

```

1      11      21      31      41      51
|      |      |      |      |      |
MIPFLPMFSL LLLLIIVNPIN ANNHYDKILA HSRIRGRDQG PNVCALQQIL GTKKKYFSTC 60
KNWYKKSICG QKTTVLYECC PGYMRMEGMR GCPAVLPIDH VYGTGLGIVGA TTTQRYSDAS 120
KLREEIEGKG SFTYFAPSNE AWDNLDSDIR RGLSENVNVE LLNALHSHMI NKRMLTKDLK 180
NGMIIPSMYN NLGLFINHYP NGVVTNINCAR IIHGNQIATN GVVHVIDRVL TQIGTSIQDF 240
IEAEDDLSSF RAAAITSDIL EALGRDGHFT LFAPTNEAFE KLPRGVLERF MGDKVASEAL 300
MKYHILNTLQ CSESIMGGAV FETLEGNTIE IGCDDGSITV NGIKMVNKKD IVTNNGVIHL 360
IDQVLIPDSA KQVIELAGKQ QTTFTDLVAQ LGLASALRPD GEYTLAPVN NAFSDDTLISM 420
VQRLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKGSQKG 480
RNGAIHIFRE IIKPAEKSLH EKLKQDKRFS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA 540
FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKFEPG VTNILKTTQG SKIFLKEVND 600
TLVNELKSK ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST 660
FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIKTEGPTL TKVKIEGEPE FRLIKEGETI 720
TEVIHGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL 780
QEEVTKVTKF IEGGDGHLFE DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ

```

Seq ID NO: 402 DNA sequence  
Nucleic Acid Accession #: NM\_002416  
Coding sequence: 40..417

65  
70  
75  
80  
85

```

1      11      21      31      41      51
|      |      |      |      |      |
ATCCAATACA GGAGTGACTT GGAACTCCAT TCTATCACTA TGAAGAAAAG TGGTGTCTTT 60
TTCTCTTGG GCATCATCTT GCTGGTTCTG ATTGGAGTGC AAGGAACCCC AGTAGTGAGA 120
AAGGGTCGCT GTTCTGCAT CAGCACCAAC CAAGGGACTA TCCACCTACA ATCCTTGAAA 180
GACCTTAAAC AATTTGCCCC AAGCCCTTCC TGCAGAAAAA TTGAAATCAT TGCTACACTG 240
AAGAATGGAG TTCAAACATG TCTAAACCCA GATTCAGCAG ATGTGAAGGA ACTGATTAAA 300
AAGTGGGAGA AACAGGTCAG CCAAAGAAA AAGCAAAAGA ATGGGAAAAA ACATCAAAAA 360
AAGAAAGTTC TGAAAGTTCT AAAATCTCAA CGTTCTCGTC AAAAGAAGAC TACATAAGAG 420
ACCACTTCAC CAATAAGTAT TCTGTGTTAA AAATGTTCTA TTTTAATTAT ACCGCTATCA 480
TTCCAAAGGA GGATGGCATA TAATACAAAG GCTTATTAAT TTGACTAGAA AATTTAAAAC 540
ATTACTCTGA AATTGTAAC AAAGTTAGAA AGTTGATTTT AAGAATCCAA ACGTTAAGAA 600
TTGTTAAAGG CTATGATTGT CTTTGTCTTT CTACCAACCA CCAGTTGAAT TTCATCATGC 660
TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT GTTCACCCAA CCACATCCCA 720
CTCACAACAG CTGCCTGGAA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCCAGAGAG 780
TATCTGAGGC ACATGTCAGC AAGTCCTAAG CCTGTAGCA TGCTGGTGAG CCAAGCAGTT 840
TGAAATTGAG CTGGACCTCA CCAAGCTGCT GTGGCCATCA ACCTCTGTAT TTGAATCAGC 900
CTACAGGCCT CACACACAAT GTGTCTGAGA GATTCATGCT GATTGTTATT GGGTATCACC 960
ACTGGAGATC ACCAGTGTGT GGCTTTCAGA GCCTCTTTC TGGCTTTGGA AGCCATGTGA 1020
TTCCATCTTG CCCGCTCAGG CTGACCACTT TATTTCTTTT TGTTCCCCTT TGCTTCATTC 1080
AAGTCAGCTC TTCTCCATCC TACCACAATG CAGTGCCTTT CTTCTCTCCA GTGCACCTGT 1140
CATATGCTCT GATTATCTG AGTCAACTCC TTTCTCATCT TGTCCCCAAC ACCCCACAGA 1200
AGTGCTTTCT TCTCCAATT CATCCTCACT CAGTCCAGCT TAGTTCAAGT CCTGCCTCTT 1260
AAATAAACCT TTTTGGACAC ACAAATTATC TTAAACTCC TGTTTCACTT GGTTCAGTAC 1320
CACATGGGTG AACACTCAAT GGTAACTAA TTCTTGGGTG TTTATCCTAT CTCTCCAACC 1380

```



AGATTGTCAG CTCCTTGAGG GCAAGAGCCA CAGTATATTT CCCTGTTTCT TCCACAGTGC 1440  
CTAATAATAC TGTGGAAC TA GGTTTTAATA ATTTTTTAAT TGATGTTGTT ATGGGCAGGA 1500  
TGGCAACCAG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCCTGGCTAC TCCATGTTGG 1560  
CTAGCCTCTG GTAACCTCTT ACTTATTATC TTCAGGACAC TCACTACAGG GACCAGGGAT 1620  
5 GATGCAACAT CCTTGTCTTT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA 1680  
GCACGTGGTA AAACACTTGC GGATATTCTG GACTGTTTTT AAAAAATATA CAGTTTACCG 1740  
AAAATCATAT AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC 1800  
CCAACCATAC AAAAATTCCT TTTCCCGAAG GAAAAGGGCT TTCTCAATAA GCCTCAGCTT 1860  
TCTAAGATCT AACAAGATAG CCACCGAGAT CCTTATCGAA ACTCATTTTA GGCAATATG 1920  
10 AGTTTTATTG TCCGTTTACT TGTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA 1980  
TCTCCCATGA AGAAAGGGAA CCGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT 2040  
TAGTGGAGC ATGATTGGTG CCCAGTTAGC CTCTGCAGGA TGTGGAAACC TCCTTCCAGG 2100  
GGAGGTTTCA TGAATTGTGT AGGAGAGGTT GTCTGTGGCC AGAATTTAAA CCTATACTCA 2160  
CTTTCCCAAA TTGAATCACT GCTCACACTG CTGATGATTT AGAGTGCTGT CCGGTGGAGA 2220  
15 TCCCACCCGA ACGTCTTATC TAATCATGAA ACTCCCTAGT TCCTTCATGT AACTTCCCTG 2280  
AAAAATCTAA GTGTTTCATA AATTTGAGAG TCTGTGACCC ACTTACCTTG CATCTCACAG 2340  
GTAGACAGTA TATAACTAAC AACCAAAGAC TACATATTGT CACTGACACA CACGTTATAA 2400  
TCAATTTATCA TATATATACA TACATGCATA CACTCTCAAA GCAAATAATT TTTCACCTCA 2460  
20 AAACAGTATT GACTTGTATA CCTTGTAATT TGAAATATTT TCTTTGTAA AATAGAATGG 2520  
TATCAATAAA TAGACCATTA ATCAG

Seq ID NO: 403 Protein sequence  
Protein Accession #: NP\_002407

1 11 21 31 41 51  
| | | | | |  
MKKSGVLFLL GIILLVLIGV QGTPVVRKGR CSCISTNQGT IHLQSLKDLK QFAPSPSCEK 60  
IEIIATLKNQ VQTCLNPDSA DVKELIKKWE KQVSQKKKQK NGKKHQKKKV LKVRKSQRSR 120  
QKKT

Seq ID NO: 404 DNA sequence  
Nucleic Acid Accession #: NM\_006670  
Coding sequence: 85..1347

1 11 21 31 41 51  
| | | | | |  
CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCAGCCC 60  
AGCTCCGGGG AAACGCGAGC CGCGATGCCT GGGGGGTGCT CCCGGGGCCC CGCCGCCGGG 120  
GACGGGCGTC TGCGGCTGGC GCGACTAGCT CTGGTACTCC TGGGCTGGGT CTCTCGTCT 180  
40 TCTCCACCTT CCTCGGCATC CTCCTTCTCC TCCTCGGCGC CGTTCCTGGC TTCCGCCGTG 240  
TCCGCCAGC CCCCCTGCGG GGACCACTGC CCGCGCTGT GCGAGTGCTC CGAGGCAGCG 300  
CGCACAGTCA AGTGCCTTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC 360  
GTGCGCAACC TCTTCTTAC CGGCAACCAG CTGGCCGTGC TCCCTGCCGG CGCCTTCGCC 420  
CGCCGGCCGC CGCTGGCGGA GCTGGCCGCG CTCAACCTCA GCGGCAGCCG CCTGGACGAG 480  
45 GTGCGGCGGG GCGCCTTCGA GCATCTGCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC 540  
CCACTGGCCG ACCTCAGTCC CTTCTGCTTT TCGGGCGACA ATGCCAGCGT CTCGGCCCCC 600  
AGTCCCCTTG TGGAACCTGAT CCTGAACCAC ATCGTGCCCC CTGAAGATGA GCGGCAGAAC 660  
CGGAGCTTCG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCGGTGCACT GCAGGGGCTC 720  
CGCCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCAA 780  
50 CTGCCCAGCC TCAGGCACCT GGAATTAAGT AATAATTCGC TGGTGAGCCT GACCTACGTG 840  
TCCTTCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900  
CTTCAACATG GCACCTGGC TGAGTTGCAA GGTCTACCCC ACATTAGGGT TTTCTTGGAC 960  
AACAATCCCT GGGTCTGCGA CTGCCACATG GCAGACATGG TGACCTGGCT CAAGGAAACA 1020  
GAGGTAGTGC AGGGCAAAGA CCGGCTCACC TGTGCATATC CGGAAAAAAT GAGGAATCGG 1080  
55 GTCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG 1140  
CAAACCTCTT ATGTCTTCTT GGGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCTCTCTG 1200  
GTTTTGTATT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCCCTG 1260  
AGGGATCACA TGGAAGGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA 1320  
AACCTCAGTT CTAACCTGGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACTCTG 1380  
60 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA 1440  
TAGATACAAC GGACTTTGAC TAAAAGCAGT GAAGGGGATT TGCTTCTTGG TTATGTAAAG 1500  
TTTCTCGGTG TGTCTGTGTA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCCTC 1560  
TTCTTTTCTT TGGAACCTCT CAACACGTAT GGAGGGATTT TTCAGGTTTC AGCATGAACA 1620  
TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680  
65 ACAGATAGCA TTCAACAAAA GCTGCTCAA CTTTTTCGAG AAAAAATACTT TATTCATAAA 1740  
TATCAGTTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAAT AATTGCATCC TATAAACTGC 1800  
CTGCAGACGT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA 1860  
AGAGCATGCT TACATTTTAC TGTCTGTCAT ATTACAAAAA ATAACCTTGA ACTTCATAAC 1920  
70 TTCTTTGACA AAGTAAATTA CTTTTTTGAT TGCAGTTTAT ATGAAAATGT ACTGATTTTT 1980  
TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAAATAAG 2040  
ATTCTTAAAA GAA

Seq ID NO: 405 Protein sequence  
Protein Accession #: NP\_006661

1 11 21 31 41 51  
| | | | | |  
MPGGCSRGP AGDGLRLRL LALVLLGWVS SSSPTSSASS FSSAPFLAS AVSAQPPLPD 60  
QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFITG NQLAVLPAGA FARRPPLAEL 120  
80 AALNLSGSRL DEVRAGAFEH LPSLRQLDLS HNPLADLSPF AFSGSNASVS APSPLVELIL 180  
NHIVPPEDER QNRSFEGMVV AALLAGRALQ GLRRLELASN HFLYLPDRDL AQLPSLRHLD 240  
LSNNSLVSLT YVSFRNLTHL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVF LDNNPWVDCD 300  
HMADMVTLWK ETEVVQKDR LTCAYPEKMR NRVLLELNSA DLDCDPIPPP SLQTSYVFLG 360  
85 IVLALIGAIF LLVLYLNRKG IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSSNSDV

Seq ID NO: 406 DNA sequence  
Nucleic Acid Accession #: Eos sequence



Coding sequence: 1..927

	1	11	21	31	41	51	
5							
	ATGCCTGGGG	GGTGTCTCCG	GGGCCCCGCC	GCCGGGGACG	GGCGTCTGCG	GCTGGCGCGA	60
	CTAGCGCTGG	TACTCCTGGG	CTGGGTCTCC	TCGTCTTCTC	CCACCTCCTC	GGCATCCTCC	120
	TTCTCCTCCT	CGGCGCCGTT	CCTGGCTTCC	GCCGTGTCCG	CCCAGCCCCC	GCTGCCGGAC	180
	CAGTGCCCCG	CGCTGTGCGA	GTGCTCCGAG	GCAGCGCGCA	CAGTCAAGTG	CGTTAACCGC	240
10	AATCTGACCG	AGGTGCCAC	GGACCTGCCC	GCCTACGTGC	GCAACCTCTT	CCTTACCGGC	300
	AACCAGCTGG	CCAGCAACCA	CTTCCTTTAC	CTGCCGCGGG	ATGTGCTGGC	CCAACTGCCC	360
	AGCCTCAGGC	ACCTGGACTT	AAGTAATAAT	TCGCTGGTGA	GCCTGACCTA	CGTGTCTTTC	420
	CGCAACCTGA	CACATCTAGA	AAGCCTCCAC	CTGGAGGACA	ATGCCCTCAA	GGTCCTTCAC	480
	AATGGCACCC	TGGCTGAGTT	GCAAGGTCTA	CCCCACATTA	GGGTTTTCCT	GGACAACAAT	540
	CCCTGGGTCT	CGGACTGCCA	CATGGCAGAC	ATGGTGACCT	GGCTCAAGGA	AACAGAGGTA	600
15	GTGCAGGGCA	AAGACCGGCT	CACCTGTGCA	TATCCGAAA	AAATGAGGAA	TCGGGTCTCT	660
	TTGGAATCA	ACAGTGCTGA	CCTGGACTGT	GACCCGATTC	TTCCCCATC	CCTGCAAACC	720
	TCTTATGTCT	TCCTGGGTAT	TGTTTTAGCC	CTGATAGCGC	CTATTTTCCT	CCTGGTTTTG	780
	TATTTGAACC	GCAAGGGGAT	AAAAAAGTGG	ATGCATAACA	TCAGAGATGC	CTGCAGGGAT	840
20	CACATGGAAG	GGTATCATTA	CAGATATGAA	ATCAATGCGG	ACCCCAGATT	AACAAACCTC	900
	AGTTCTAACT	CGGATGTCCT	CGAGTGA				

Seq ID NO: 407 Protein sequence  
Protein Accession #: Eos sequence

25	1	11	21	31	41	51	
	MPGGSRRGPA	AGDRLRLRL	LALVLLGWVS	SSSPTSSASS	FSSSAPFLAS	AVSAQPPLPD	60
	QCPALCECSE	AARTVKCVNR	NLTEVPTDLP	AYVRNLFLLT	NQLASNHFLY	LPRDVLAQLP	120
30	SLRHLDLN	SLVSLTYVSF	RNLTHLES	LEDNALKVLH	NGTLAELQGL	PHIRVFLDNN	180
	PWVCDCHMAD	MVTWLKETEV	VQKDRILTCA	YPEKMRNRVL	LELNSADLDC	DPILPPSLQT	240
	SYVFLGLVLA	LIGAIFLLVL	YLNKRGIKKW	MHNIRDACRD	HMEGYHYRYE	INADPRLTNL	300
	SSNSDVLE						

Seq ID NO: 408 DNA sequence  
Nucleic Acid Accession #: NM\_000095.1  
Coding sequence: 26..2299

	1	11	21	31	41	51	
40	CAGCACCCAG	CTCCCCGCCA	CCGCCATGGT	CCCCGACACC	GCCTGCGTTC	TTCTGCTCAC	60
	CCTGGCTGCC	CTCGGCGCGT	CCGGACAGGG	CCAGAGCCCG	TTGGGCTCAG	ACCTGGGCCC	120
	GCAGATGCTT	CGGGAAGTGC	AGGAAACCAA	CGCGGCGCTG	CAGGACGTGC	GGGACTGGCT	180
	GCGGCAGCAG	GTCAGGGAGA	TCACGTTTCT	GAAAAACACG	GTGATGGAGT	GTGACGCGTG	240
45	CGGGATGCAG	CAGTCAGTAC	GCACCGGCCT	ACCCAGCGTG	CGGCCCCCTG	TCCACTGCGC	300
	GCCCCGCTTC	TGCTTCCCCG	GCGTGGCCTG	CATCCAGACG	GAGAGCGGCG	GCCGCTGCGG	360
	CCCCTGCCCC	GCGGGCTTCA	CGGGCAACGG	CTCGCACTGC	ACCGACGTCA	ACGAGTGCAA	420
	CGCCCAACCC	TGCTTCCCCC	GAGTCCGCTG	TATCAACACC	AGCCCGGGGT	TCCGCTGCGA	480
	GGCTTGCCCG	CCGGGGTACA	GCGGCCCCAC	CCACCAGGGC	GTGGGGCTGG	CTTTCGCCAA	540
50	GGCCAACAAG	CAGGTTTGCA	CGGACATCAA	CGAGTGTGAG	ACCGGGCAAC	ATAACTGCGT	600
	CCCCAACTCC	GTGTGCATCA	ACACCCGGGG	CTCCTTCCAG	TGCGGCCCCG	GCCAGCCCGG	660
	CTTCGTGGGC	GACCAGGCGT	CCGGCTGCCA	GCAGCGCGCA	CAGCGCTTCT	GCCCCGACGG	720
	CTCGCCACG	GAGTGCCACG	AGCATGCAGA	CTGCGTCCTA	GAGCGCGATG	GCTCGCGGTC	780
	GTGCGTGTGT	CGCGTTGGCT	GGGCCGGCAA	CGGGATCCTC	TGTGGTGGCG	ACACTGACCT	840
55	AGACGGCTTC	CCGGACGAGA	AGCTGCGCTG	CCCGGAGCCG	CAGTGCCGTA	AGGACAACCTG	900
	CGTGACTGTG	CCCAACTCAG	GGCAGGAGGA	TGTGGACCCG	GATGGCATCG	GAGACGCCTG	960
	CGATCCGGAT	GCCGACGGGG	ACGGGGTCCC	CAATGAAAAG	GACAACTGCC	CGCTGGTGCG	1020
	GAACCCAGAC	CAGCGCAACA	CGGACGAGGA	CAAGTGGGGC	GATGCGTGCG	ACAACCTGCCG	1080
	GTCCCAAGAG	AACGACGACC	AAAAGGACAC	AGACAGGAC	GGCCGGGGCG	ATGCGTGCGA	1140
60	CGACGACATC	GACGGCGACC	GGATCCGCAA	CCAGGCCGAC	AACTGCCCTA	GGGTACCCAA	1200
	CTCAGACCAG	AAGGACAGTG	ATGGCGATGG	TATAGGGGAT	GCCTGTGACA	ACTGTCCCCA	1260
	GAAGAGCAAC	CCGGATCAGG	CGGATGTGGA	CCACGACTTT	GTGGGAGATG	CTTGTGACAG	1320
	CGATCAAGAC	CAGGATGGAG	ACGGACATCA	GGACTCTCGG	GACAACTGTC	CCACGGTGCC	1380
	TAACAGTGCC	CAGGAGGACT	CAGACCACGA	TGGCCAGGGT	GATGCCTGCG	ACGACGACGA	1440
65	CGACAAATGAC	GGAGTCCCTG	ACAGTCGGGA	CAACTGCCGC	CTGGTGCCCTA	ACCCCGGCCA	1500
	GGAGGACGCG	GACAGGGACG	GCGTGGGCGA	CGTGTGCCAG	GACGACTTTG	ATGCAGACAA	1560
	GGTGGTAGAC	AAGATCGACG	TGTGTCCGGA	GAACGCTGAA	GTCACGCTCA	CCGACTTCAG	1620
	GGCCTTCCAG	ACAGTCGTGC	TGGACCCGGA	GGGTGACGCG	CAGATTGACC	CCAACCTGGGT	1680
	GGTGCTCAAC	CAGGGAAGGG	AGATCGTGCA	GACAATGAAC	AGCGACCCAG	GCCTGGCTGT	1740
70	GGGTTACACT	GCCTTCAATG	GCGTGGACTT	CGAGGGCACG	TTCCATGTGA	ACACGGTCAC	1800
	GGATGACGAC	TATGCGGGCT	TCATCTTTTG	CTACCAGGAC	AGCTCCAGCT	TCTACGTGGT	1860
	CATGTGGAAG	CAGATGGAGC	AAACGTATTG	CGAGGCGAAC	CCCTTCCGTG	CTGTGGCCGA	1920
	GCCTGGCATC	CAACTCAAGG	CTGTGAAGTC	TTCCACAGGC	CCCGGGGAAC	AGCTGCGGAA	1980
	CGCTCTGTGG	CATACAGGAG	ACACAGAGTC	CCAGGTGCGG	CTGCTGTGGA	AGGACCCGCG	2040
75	AAACGTGGGT	TGGAAGGACA	AGAAGTCCTA	TCGTTGGTTC	CTGCAGCACC	GGCCCCAAGT	2100
	GGGCTACATC	AGGGTGCGAT	TCTATGAGGG	CCCTGAGCTG	GTGGCCGACA	GCAACGTGGT	2160
	CTTGGACACA	ACCATGCGGG	GTGGCCGCCT	GGGGGTCTTC	TGCTTCTCCC	AGGAGAACAT	2220
	CATCTGGGCC	AACCTGCGTT	ACCGCTGCAA	TGACACCATC	CCAGAGGACT	ATGAGACCCA	2280
	TCAGCTGCCG	CAAGCCTAGG	GACCAAGGTG	AGGACCCGCC	GGATGACAGC	CACCCCTACC	2340
80	GCGGCTGGAT	GGGGGCTCTG	CACCCAGCCC	AAGGGGTGGC	CGTCCTGAGG	GGGAAGTGAG	2400
	AAGGGCTCAG	AGAGGACAAA	ATAAAGTGTG	TGTGCAGGG			

Seq ID NO: 409 Protein sequence  
Protein Accession #: NP\_000086.1

85	1	11	21	31	41	51	
	MVPDTACVLL	LTLAALGASG	QGQSPLGSDL	GPQMLRELQE	TNAALQDVRD	WLRQQVREIT	60



FLKNTVMECD ACGMQQSVRT GLPSVRPLLH CAPGFCFPGV ACIQTESGGR CGPCPAGFTG 120  
NGSHCTDVNE CNAHPCFPRV RCINTSPGFR CEACPPGYSG PTHQGVGLAF AKANKQVCTD 180  
INECETGQHN CVPNSVCINT RGSFQCGPCQ PGFVGDAQSG CQRGAQRFCP DGSPSECHEH 240  
5 ADCVLERDGS RSCVCRVGWA GNGILCGRDT DLDGFPDEKL RCPEPQCRKD NCVTVPNSGQ 300  
EDVDRDGDG ACDFDADGDG VPNEKDNCPL VRNPDQRNTD EDKWDGACDN CRSQKNDDQK 360  
DTDQDGRGDA CDDDDIDGDRI RNQADNCPRV PNSDQKDSG DGIGDADNC PQKSNPDQAD 420  
VDHDFVGDAC DSDQDQDGDG HQDSRDNCPT VPNSAQEDSD HDGQGDACDD DDDNDGVVDS 480  
RDNCRLVFPN GQEDADRDGV GDVCQDDFDA DKVVDKIDVC PENAEVTLTD FRAFQTVVLD 540  
10 PEGDAQIDPN WVVLNQGREI VQTMNSDPGL AVGYTAFNGV DFEGTFHVNT VTDDDYAGFI 600  
FGYQDSSSFY VVMWKQMEQT YWQANPFRAV AEPGIQLKAV KSSTGPGEQL RNALWHTGDT 660  
ESQVRLWWD PRNVGWKDKK SYRWFLQHRP QVGYIRVRFY EGPELVADSN VVLDTTMRGG 720  
RLGVFCFSQE NIIWANLRYR CNDTIPEDYE THQLRQA

Seq ID NO: 410 DNA sequence  
Nucleic Acid Accession #: NM\_001565.1  
Coding sequence: 67..363

1 11 21 31 41 51  
| | | | | |  
20 GAGACATTCC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC CTCCAGTCTC 60  
AGCACCATGA ATCAAACTGC GATTCTGATT TGCTGCCCTA TCTTTCTGAC TCTAAGTGGC 120  
ATTCAAGGAG TACCTCTCTC TAGAACCGTA CGCTGTACCT GCATCAGCAT TAGTAATCAA 180  
CCTGTTAATC CAAGGTCTTT AGAAAACTT GAAATTATTC CTGCAAGCCA ATTTTGTCCA 240  
25 CGTGTGAGA TCATTGCTAC AATGAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA 300  
TCGAAGGCCA TCAAGAATTT ACTGAAAGCA GTTAGCAAGG AAATGTCTAA AAGATCTCCT 360  
TAAAACCCAG GGGGAGCAAA ATCGATGCAG TGCTTCCAAG GATGGACCAC ACAGAGGCTG 420  
CCTCTCCCAT CACTTCCCTA CATGGAGTAT ATGTCAAGCC ATAATTGTTT TTAGTTTGCA 480  
GTTACACTAA AAGGTGACCA ATGATGGTCA CCAATCAGC TGCTACTACT CCTGTAGGAA 540  
GGTTAATGTT CATCATCCTA AGCTATTGAG TAATAACTCT ACCCTGGCAC TATAATGTAA 600  
30 GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTTCTGACC CTGCTTCAA TATTTCCCTC 660  
ACCTTTCCCA TCTTCCAAGG GTACTAAGGA ATCTTTCTGC TTTGGGGTTT ATCAGAATTC 720  
TCAGAATCTC AAATAACTAA AAGGTATGCA ATCAAATCTG CTTTTTAAAG AATGCTCTTT 780  
ACTTCATGGA CTTCCACTGC CATCCTCCCA AGGGGCCCAA ATTCTTTTCTG TGGCTACCTA 840  
35 CATACAATTC CAAACACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAAGTATT 900  
CTTATTTAAT GAAAGACTGT ACAAAGTATA AGTCTTAGAT GTATATATTT CCTATATTGT 960  
TTTCAGTGTA CATGGAATAA CATGTAATTA AGTACTATGT ATCAATGAGT AACAGGAAAA 1020  
TTTTAAAAAT ACAGATAGAT ATATGCTCTG CATGTTACAT AAGATAAATG TGCTGAATGG 1080  
TTTTCAAATA AAAATGAGGT ACTCTCCTGG AAATATTAAG

Seq ID NO: 411 Protein sequence  
Protein Accession #: NP\_001556.1

1 11 21 31 41 51  
| | | | | |  
45 MNQTAILICC LIFLTLSGIQ GVPLSRTVRC TCISISNQPV NPRSLEKLEI IPASQFCPRV 60  
EIIATMKKKG EKRLNPESK AIKNLLKAVS KEMSKRSP

Seq ID NO: 412 DNA sequence  
Nucleic Acid Accession #: XM\_057014  
Coding sequence: 143..874

1 11 21 31 41 51  
| | | | | |  
55 GGGAGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCCTCGGAG 60  
CGCGGCGGAG CCAGACGCTG ACCACGTTCC TCTCCTCGGT CTCCTCCGCC TCCAGCTCCG 120  
CGCTGCCCCG CAGCCGGGAG CCATGCGACC CCAGGGCCCC GCGCCTCCC CGCAGCGGCT 180  
CCGCGGCCCTC CTGCTGCTCC TGCTGCTGCA GCTGCCCGCG CCGTCGAGCG CCTCTGAGAT 240  
60 CCCCAGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG 300  
AATGTGCTTA CAAGGGCCAG CAGGAGTGCC TGGTCGAGAC GGGAGCCCTG GGGCCAATGG 360  
CATTCGCGGT ACACCTGGGA TCCCAGGTCG GGATGGATTC AAAGGAGAAA AGGGGGAATG 420  
TCTGAGGGAA AGCTTTGAGG AGTCCTGGAC ACCCAACTAC AAGCAGTGTT CATGGAGTTC 480  
ATTGAATTAT GGCATAGATC TTGGGAAAAA TGCGGAGTGT ACATTTACAA AGATGCGTTC 540  
AAATAGTGCT CTAAGAGTTT TGTTCACTGG CTCACCTTCG CTAAAATGCA GAAATGCATG 600  
CTGTCAGCGT TGGTATTTCA CATTCAATGG AGCTGAATGT TCAGGACCTC TTCCCATTTGA 660  
65 AGCTATAATT TATTTGGACC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTCATCG 720  
CACTTCTTCT GTGGAAGGAC TTTGTGAAGG AATTGGTGCT GGATTAGTGG ATGTTGCTAT 780  
CTGGGTTGCG ACTTGTTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT 840  
TTCTCGCATC ATTATTGAAG AACTACCAA ATAAATGCTT TAATTTTCAT TTGCTACCTC 900  
TTTTTTTATT ATGCCTTGGA ATGGTTCACT TAAATGACAT TTAAATAAG TTTATGTATA 960  
70 CATCTGAATG AAAAGCAAAG CTAAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT 1020  
TTTAAATCTA GCATTATTCA TTTTGCTTCA ATCAAAAGTG GTTTCAATAT TTTTCTTAGT 1080  
TGGTTAGAA ACTTCTTCTA TAGTCACATT CTCTCAACCT ATAATTTGGA ATATTGTTGT 1140  
GGTCTTTTGT TTTTCTCTT AGTATAGCAT TTTTAAAAAA ATATAAAAGC TACCAATCTT 1200  
75 TGTACAATTT GTAAATGTTA AGAATTTTTT TTATATCTGT TAAATAAAAA TTATTTCCAA 1260  
CAACCTTAAA AAAAAAAAAA AAAA

Seq ID NO: 413 Protein sequence  
Protein Accession #: XP\_057014

1 11 21 31 41 51  
| | | | | |  
80 MRPQGPAAASP QRLRGLLLLL LLQLPAPSSA SEIPKQKQKA QLRQREVVDL YNGMCLQGPA 60  
GVPGRDGSFG ANGIPGTPGI PGRDGFKEG GECLRESFEE SWTPNYKQCS WSSLNYGIDL 120  
85 GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACCQRWYFT FNGAECGSL PLIEAIIYLDQ 180  
GSPEMNSTIN IHRTSSVEGL CEGIGAGLVD VAIWVGTCSD YPKGDASTGW NSVSRIIEE 240  
LPK



WO 02/086443

Seq ID NO: 414 DNA sequence  
Nucleic Acid Accession #: XM\_084007  
Coding sequence: 138..2405

PCT/US02/12476

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      CTCGTGCCGA ATTCGGCAGC AGACCGCGTG TTCGCGCCTG GTAGAGATTT CTCGAAGACA      60
      CCAGTGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCCG GGCCGTGGGA CAACGAGGCC      120
10     GCGGAGACGA AGGCGCAATG GCGAGGAAGT TATCTGTAAT CTTGATCCTG ACCTTTGCCC      180
      TCTCTGTAC AAATCCCCTT CATGAACATA AAGCAGCTGC TTTCCCCCAG ACCACTGAGA      240
      AAATTAGTCC GAATTGGGAA TCTGGCATTG ATGTTGACTT GGCAATTTCC ACACGGCAAT      300
      ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTGTCA GTTGAAGGGT      360
      TCAGAAAATT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG      420
15     ACCACGACCA TCACTCAGAC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG      480
      AGCATCACTC AGACCACGAG CATCACTCTG ACCATGATCA TCACTCCAC CATAATCATG      540
      CTGCTTCTGG TAAAAATAAG CGAAAAGCTC TTTGCCCAGA CCATGACTCA GATAGTTCAG      600
      GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA      660
      GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCAACTGTG TACAACACTG      720
20     TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACCTCTCC      780
      CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTACATC AAAGAGCCGG GTGAGCCGGC      840
      TGGCTGGTAG GAAAACAAAT GAATCTGTGA GTGAGCCCCG AAAAGGCTTT ATGTATTCCA      900
      GAAACACAAA TGAAAATCCT CAGGAGTGTT TCAATGCATC AAAGCTACTG ACATCTCATG      960
      GCATGGGCAT CCAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA      1020
25     TCAACCAAAT TGATGCTAGA TCTTGCTCTG TTCATACAAG TGAAAAGAAG GCTGAAATCC      1080
      CTCCAAAGAC CTATTCATTA CAAATAGCCT GGGTTGGTGG TTTTATAGCC ATTTCCATCA      1140
      TCAGTTTCTT GTCTCTGCTG TGGGTTATCT TAGTGCCTCT CATGAATCGG GTGTTTTTCA      1200
      AATTTCTCCT GAGTTTCTTT GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTTT      1260
      TACACCTTCT TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG      1320
30     CAATGGAAT GAAAAGAGGA CCACTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA      1380
      GTGCCTATTT TGATTCCACG TGAAGGGTTC TAACAGCTCT AGGAGGCCTG TATTTTCATG      1440
      TTCTTGTTGA ACATGTCCTC ACATTGATCA AACAATTTAA AGATAAGAAG AAAAAGAATC      1500
      AGAAGAAACC TGAAAATGAT GATGATGTGG AGATTAAAGAA GCAGTTGTCC AAGTATGAAT      1560
35     CTCAACTTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTTAC      1620
      GAGCAGACTC ACAAGAGCCC TCCCACCTTTG ATTCTCAGCA GCCTGCAGTC TTGGAAGAAG      1680
      AAGAGGTCAT GATAGCTCAT GCTCATCCAC AGGAAGCTCA CAATGAATAT GTACCCAGAG      1740
      GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA      1800
      TTCACCACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCACCA AACCACCATC      1860
      CTCACAGTCA CAGCCAGCGC TACTCTCGGG AGGAGCTGAA AGATGCCGGC GTCGCCACTT      1920
40     TGGCCTGGAT GGTGATAATG GGTGATGGCC TGCACAATTT CAGCGATGGC CTAGCAATTG      1980
      GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTTGCT GTGTTCTGTC      2040
      ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA      2100
      AGCAGGCTGT CCTTTATAAT GCATTGTGAG CCATGCTGGC GTATCTTGGG ATGGCAACAG      2160
      GAATTTTCAT TGGTCATTAT GCTGAAAATG TTTCTATGTG GATATTTGCA CTTACTGCTG      2220
45     GCTTATTCAT GTATGTTGCT CTGGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA      2280
      GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTACA GAATGCTGGG ATGCTTTTGG      2340
      GTTTTGGAAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATCGTGTTT CGTATAAATT      2400
      TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAAGTTG TCATAGTTTC AGTAGGTCAT      2460
      AGGGAGATGA GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGGT TTTGTGATTT      2520
50     TTGTATTGAA TATTGCTGTC TGTTACAAAG TCAGTTAAAG GTACGTTTAA ATATTAAAGT      2580
      TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG      2640
      TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTCA GGAATAATGT CTTAATGCT      2700
      TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACTGCTGG      2760
      TGTTTAGGAA TAAGAATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA      2820
55     AGCAAAGAAA TAAAGGAGAA AAGAGAAGAA TCTGAGAAAT GGGGAGGCAT AGATTCTTAT      2880
      AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAAG TATAAAAAGG      2940
      CAGAATTAGT ATAGAGTACA TTCATTAAC ATTTTGTGCA GGATTATTTT CCGTAAAAAC      3000
      GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTTAAC TTTGTATAAT ACAGAAATCT      3060
      AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTGGAA TTTCAAAATG      3120
      TTCGTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT      3180
60     TATTGCCAAG TTATATATCA CCAAAGCTG TATGACTGGA TGTTCTGGTT ACCTGGTTTA      3240
      CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAACT ACCTAAGTA      3300
      TCATTTGATT CGATTGAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTGT      3360
      GAGCAATTGT CTTTATATAC GGTACTGTAG CCATACTAGG CCTGCTGTG GCATTCCTTA      3420
65     GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCTT G

```

Seq ID NO: 415 Protein sequence  
Protein Accession #: XP\_084007

```

70     1      11      21      31      41      51
      |      |      |      |      |      |
      MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPNN ESGINVDLAI STRQYHLQQL      60
      FYRYGENNSL SVEGFRKLLQ NIGIDKIKRI HIHHDHHDHS DHEHSDHER HSDHEHSDH      120
      EHHSDHDHHS HHNHAASGKN KRKALCPDHD SDSSGKDPNN SQGKAHRPE HASGRNVKD      180
75     SVSASEVTST VYNTVSEGTH FLETIETPRP GKLPKDVSS STPPSVTSKS RVSRLAGRKT      240
      NESVSEPRKG FMYSRNTNEN PQECFVNSGL LSHGMIQV PLNATEFNYL CPAINQIDA      300
      RSLIHTSEK KAEIPPKTYS LQIAWVGGFI AISIISFLSL LGVILVPLMN RVFFKFLLSF      360
      LVALAVGTLS GDAFLHLLPH SHASHHSHS HEPPAMEMKR GPLFSLSSQ NIEESAYFDS      420
      TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKNQKKPEN DDDVEIKKQL SKYESQLSTN      480
80     EEKVDTDRT EGYLRADSQE PSHFDSQQA VLEEEVMIH HAHPQEVYNE YVPRGCKNKC      540
      HSHFHDTLGQ SDDLIHHHHD YHHILHHHSH QNHHPHSHSQ RYSREELKDA GVATLAWMVI      600
      MGDGLHNFSD GLAIGAAFTG GLSSGLSTSV AVFCHLPHE LGDFAVLLKA GMTVKQAVLY      660
      NALSAMLAYL GMATGIFIGH YAENVSMIV ALTAGLFMYV ALVDMVPEML HNDASDHGCS      720
      RWGYFFLQNA GMLLGFGIML LISIFEHKIV FRINF

```

Seq ID NO: 416 DNA sequence  
Nucleic Acid Accession #: NM\_015419.1  
Coding sequence: 1..8487



	1	11	21	31	41	51	
5	ATGCCCAAGC	GCGCGCACTG	GGGGGCCCTC	TCCGTGGTGC	TGATCCTGCT	TTGGGGCCAT	60
	CCGCGAGTGG	CGCTGGCCTG	CCCGCATCCT	TGTGCCTGCT	ACGTCCCCAG	CGAGGTCCAC	120
	TGCACGTTCC	GATCCCTGGC	TTCCGTGCCC	GCTGGCATTG	CTAGACACGT	GGAAAGAATC	180
	AATTTGGGGT	TTAATAGCAT	ACAGGCCCTG	TCAGAAACCT	CATTTGCAGG	ACTGACCAAG	240
	TTGGAGCTAC	TTATGATTCA	CGGCAATGAG	ATCCCAAGCA	TCCCCGATGG	AGCTTTAAGA	300
10	GACCTCAGCT	CTCTTCAGGT	TTTCAAGTTC	AGCTACAACA	AGCTGAGAGT	GATCACAGGA	360
	CAGACCTCC	AGGGTCTCTC	TAACTTAATG	AGGCTGCACA	TTGACCACAA	CAAGATCGAG	420
	TTTATCCACC	CTCAAGCTTT	CAACGGCTTA	ACGTCTCTGA	GGCTACTCCA	TTTGGAAGGA	480
	AATCTCCTCC	ACCAGCTGCA	CCCCAGCACC	TTCTCCACGT	TCACATTTTT	GGATTATTTT	540
	AGACTCTCCA	CCATAAGGCA	CCTCTACTTA	GCAGAGAACA	TGGTTAGAAC	TCTTCCTGCC	600
15	AGCATGCTTC	GGAACATGCC	GCTTCTGGAG	AATCTTTACT	TGCAGGGAAA	TCCGTGGACC	660
	TGCGATTGTG	AGATGAGATG	GTTTGTGGAA	TGGGATGCAA	AATCCAGAGG	AATTCTGAAG	720
	TGTAAAAAGG	ACAAAGCTTA	TGAAGGCGGT	CAGTTGTGTG	CAATGTGCTT	CAGTCCAAAG	780
	AAGTTGTACA	AACATGAGAT	ACACAAGCTG	AAGGACATGA	CTTGTCTGAA	GCCTTCAATA	840
	GAGTCCCCTC	TGAGACAGAA	CAGGAGCAGG	AGTATTGAGG	AGGAGCAAGA	ACAGGAAGAG	900
20	GATGGTGGCA	GCCAGCTCAT	CCTGGAGAAA	TTCCAACTGC	CCCAGTGGAG	CATCTCTTTG	960
	AATATGACCG	ACGAGCACGG	GAACATGGTG	AACTTGGTCT	GTGACATCAA	GAAACCAATG	1020
	GATGTGTACA	AGATTCACCT	GAACCAAAAC	GATCCTCCAG	ATATTGACAT	AAATGCAACA	1080
	GTTGCCCTTG	ACTTTGAGTG	TCCAATGACC	CGAGAAAAC	ATGAAAAGCT	ATGGAATTTG	1140
	ATAGCATACT	ACAGTGAAGT	TCCCGTGAAG	CTACACAGAG	AGCTCATGCT	CAGCAAAGAC	1200
25	CCCAGAGTCA	GCTACCACTA	CAGGCAGGAT	GCTGATGAGG	AAGCTCTTTA	CTACACAGGT	1260
	GTGAGAGCCC	AGATTCTTGC	AGAACCAGAA	TGGGTCTATG	AGCCATCCAT	AGATATCCAG	1320
	CTGAACCGAC	GTCAGAGTAC	GGCCAAGAAG	GTGCTACTTT	CCTACTACAC	CCAGTATTCT	1380
	CAACAATAT	CCACCAAAGA	TACAAGGCAG	GCTCGGGGCA	GAAGCTGGGT	AATGATTGAG	1440
	CCTAGTGGAG	CTGTGCAAAG	AGATCAGACT	GTCTTGGGAA	GGGGTCCATG	CCAGTTGAGC	1500
30	TGCAACGTGA	AAGCTTCTGA	GAGTCCATCT	ATCTTCTGGG	TGCTTCCAGA	TGGCTCCATC	1560
	CTGAAAGCGC	CCATGGATGA	CCCAGACAGC	AAGTTCTCCA	TTCTCAGCAG	TGGCTGGCTG	1620
	AGGATCAAGT	CCATGGAGCC	ATCTGACTCA	GGCTTGTACC	AGTGCATTGC	TCAAGTGAGG	1680
	GATGAAATGG	ACCGCATGGT	ATATAGGGTA	CTTGTGCAGT	CTCCCTCCAC	TCAGCCAGCC	1740
	GAGAAAGACA	CAGTGACAAT	TGGCAAGAAC	CCAGGGGAGT	CGGTGACATT	GCCTTGCAAT	1800
35	GCTTTAGCAA	TACCCGAAGC	CCACTTATAG	TGGATTCTTC	CAAACAGAAG	GATAATTAAT	1860
	GATTTGGCTA	ACACATCACA	TGTATACATG	TTGCCAAATG	GAATCTTTTC	CATCCCAAAG	1920
	GTCCAAGTCA	GTGATAGTGG	TTACTACAGA	TGTGTGGCTG	TCAACCAGCA	AGGGGCAGAC	1980
	CATTTTACGG	TGGGAATCAC	AGTGACCAAG	AAAGGGTCTG	GCTTGCCATC	CAAAAGAGGC	2040
	AGACGCCCCG	GTGCAAAGGC	TCTTTCCAGA	GTCAGAGAAG	ACATCGTGGA	GGATGAAGGG	2100
40	GGCTCGGGCA	TGGGAGATGA	AGAGAACACT	TCAAGGAGAC	TTCTGCATCC	AAAGGACCAA	2160
	GAGGTGTTCC	TCAAAACAAA	GGATGTGCC	ATCAATGGAG	ACAAGAAAGC	CAAGAAAGGG	2220
	AGAAGAAAGC	TGAAACTCTG	GAAGCATTCG	GAAAAAGAAC	CAGAGACCAA	TGTTGCAGAA	2280
	GGTCGCAGAG	TGTTTGAATC	TAGACGAAGG	ATAAACATGG	CAAACAAACA	GATTAATCCG	2340
	GAGCGCTGGG	CTGATATTTT	AGCCAAAGTC	CGTGGGAAAA	ATCTCCCTAA	GGGCACAGAA	2400
45	GTACCCCCGT	TGATTAATAAC	CACAAGTCCT	CCATCCTTGA	GCCTAGAAGT	CACACCACCT	2460
	TTTCTGCTG	TTTCTCCCCC	CTCAGCATCT	CCTGTGCAGA	CAGTAACCAG	TGCTGAAGAA	2520
	TCCTCAGCAG	ATGTACCTCT	ACTTGGTGAA	GAAGAGCACG	TTTTGGGTAC	CATTTCTCTA	2580
	GCCAGCATGG	GGCTAGAACA	CAACCACAAT	GGAGTTATTC	TTGTTGAACC	TGAAGTAACA	2640
	AGCACACCTC	TGGAGGAAGT	TGTTGATGAC	CTTTCTGAGA	AGACTGAGGA	GATAACTTCC	2700
50	ACTGAAGGAG	ACCTGAAGGG	GACAGCAGCC	CCTACACTTA	TATCTGAGCC	TTATGAACCA	2760
	TCTCCTACTC	TGCACACATT	AGACACAGTC	TATGAAAAGC	CCACCCATGA	AGAGACGGCA	2820
	ACAGAGGGTT	GGTCTGCAGC	AGATGTTGGA	TCGTCACCAG	AGCCACATC	CAGTGAGTAT	2880
	GAGCCTCCAT	TGGATGCTGT	CTCCTTGGCT	GAGTCTGAGC	CCATGCAATA	CTTTGACCCA	2940
	GATTTGGAGA	CTAAGTCACA	ACCAGATGAG	GATAAGATGA	AAGAAGACAC	CTTTGCACAC	3000
55	CTTACTCCAA	CCCCCACCAT	CTGGGTAAAT	GACTCCAGTA	CATCACAGTT	ATTTGAGGAT	3060
	TCTACTATAG	GGGAACCAGG	TGTCCCAGGC	CAATCACATC	TACAAGGACT	GACAGACAAC	3120
	ATCCACCTTG	TGAAAAGTAG	TCTAAGCACT	CAAGACACCT	TACTGATTAA	AAAGGGTATG	3180
	AAAGAGATGT	CTCAGACACT	ACAGGGAGGA	AATATGCTAG	AGGGAGACCC	CACACACTCC	3240
	AGAAGTTCTG	AGAGTGAGGG	CCAAGAGAGC	AAATCCATCA	CTTTGCCTGA	CTCCACACTG	3300
60	GGTATAATGA	GCAGTATGTC	TCCAGTTAAG	AAGCCTGCGG	AAACCACAGT	TGGTACCCTC	3360
	CTAGACAAAG	ACACCACAAC	AGTAACAACA	ACACCAAGGC	AAAAAGTTGC	TCCGTCTATC	3420
	ACCATGAGCA	CTCACCCTTC	TGGAAGGAGA	CCCAACGGGA	GAAGGAGATT	ACGCCCCAAC	3480
	AAATTCGCGC	ACCGGCACAA	GCAAACCCCA	CCCACAACCT	TTGCCCCATC	AGAGACTTTT	3540
	TCTACTCAAC	CAACTCAAGC	ACCTGACATT	AAGATTTCAA	GTCAAGTGGA	GAGTTCTCTG	3600
65	GTTCTTACAG	CTTGGGTGGA	TAACACAGTT	AATACCCCA	AACAGTTGGA	AATGGAGAAG	3660
	AATGCAGAAC	CCACATCCAA	GGGAACACCA	CGGAGAAAAC	ACGGGAAGAG	GCCAAACAAA	3720
	CATCGATATA	CCCCCTTCTA	AGTGAGCTCA	AGAGCGTCCG	GATCCAAGCC	CAGCCCTTCT	3780
	CCAGAAAATA	AACATAGAAA	CATTGTTACT	CCCAGTTCAG	AACTATACT	TTTGCTTAGA	3840
	ACTGTTTCTC	TGAAAACCTG	GGGCCCTTAT	GATTCCCTAG	ATTACATGAC	AACCACCAGA	3900
70	AAAATATATT	CATCTTACCC	TAAAGTCCAA	GAGACACTTC	CAGTCACATA	TAAACCCACA	3960
	TCAGATGGAA	AAGAAATTAA	GGATGATGTT	GCCACAAATG	TTGACAAACA	TAAAAGTGAC	4020
	ATTTTAGTCA	CTGGTGAATC	AATTACTAAT	GCCATACCAA	CTTCTCGCTC	CTTGGTCTCC	4080
	ACTATGGGAG	AATTTAAGGA	AGAATCCTCT	CCTGTAGGCT	TTCCAGGAAC	TCCAACCTGG	4140
	AATCCCTCAA	GGACGGCCCA	GCCTGGGAGG	CTACAGACAG	ACATACCTGT	TACCACTTCT	4200
75	GGGGAAAATC	TTACAGACCC	TCCCCTTCTT	AAAGAGCTTG	AGGATGTGGA	TTTCACTTCC	4260
	GAGTTTTTGT	CCTCTTTGAC	AGTCTCCACA	CCATTTTACC	AGGAAGAAGC	TGGTTCTTCC	4320
	ACAACCTCTC	CAAGCATAAA	AGTGGAGGTG	GCTTCAAGTC	AGGCAGAAAC	CACCACCCTT	4380
	GATCAAGATC	ATCTTGAAC	CACTGTGGCT	ATTCTCCTTT	CTGAAACTAG	ACCACAGAAT	4440
	CACACCCCTA	CTGCTGCCCG	GATGAAGGAG	CCAGCATCCT	CGTCCCCATC	CACAATTCTC	4500
80	ATGTCTTTGG	GACAAACCAC	CACCACTAAG	CCAGCACTTC	CCAGTCCAAG	AATATCTCAA	4560
	GCATCTAGAG	ATTCCAAGGA	AAATGTTTTT	TTGAATTATG	TGGGGAATCC	AGAAACAGAA	4620
	GCAACCCAG	TCAACAATGA	AGGAACACAG	CATATGTCAG	GGCCAAATGA	ATTATCAACA	4680
	CCCTCTTCCG	ACCGGGATGC	ATTTAACTTG	TCTACAAAGC	TGGAATTGGA	AAAGCAAGTA	4740
	TTTGGTAGTA	GGAGTCTACC	ACGTGGCCCA	GATAGCCCA	GCCAGGATGG	AAGAGTTCAT	4800
85	GCTTCTCATC	AACATAACCAG	AGTCCCTGCC	AAACCCATCC	TACCAACAGC	AACAGTGAGG	4860
	CTACCTGAAA	TGTCCACACA	AAGCGCTTCC	AGATACTTTG	TAACTTCCCA	GTCACCTCGT	4920
	CACTGGACCA	ACAAACCGGA	AATAACTACA	TATCCTTCTG	GGGCTTTGCC	AGAGAACAAA	4980
	CAGTTTACAA	CTCCAAGATT	ATCAAGTACA	ACAATTCCTC	TCCCATTGCA	CATGTCCAAA	5040



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80

CCCAGCATTC CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTACTCCAAA 5100  
GTGTTTGGAA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC TCCCAGTCCA 5160  
AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA CCAACAAGAC TCTTTCTTTT 5220  
CCACAGTTGG GAGTACCCCG GAGACCCCG ATACCCACTT CTCTGCCCC AGTAATGAGA 5280  
GAGAGAAAAG TTATTCCAGG TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG 5340  
GACTTTGGCC CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCCTCA 5400  
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC CTTTATAACA 5460  
TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT CAAAGTTCTT TGCAGGAGGA 5520  
CCTCCTGCAT CCAAATTCTG GTCTCTTGGG GAAAAGCCCC AAATCCTCAC CAAGTCCCCA 5580  
CAGACTGTGT CCGTCACCGC TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA 5640  
CCAAAGCCTT TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC 5700  
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA GGTTCAAGTA 5760  
CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC ACGGCCTGGA CAGGATGGTG 5820  
GTCTTGCTTT CGGTACCGT GCAGCAACCT CAAATCCTAG CCTCCACTA CCAGGACGTC 5880  
ACTGTCTACC TGGGAGACAC CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC 5940  
CAAATTTCTT GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC CGTGGAGAGC 6000  
CGCATCACCC TGCACGAAAA CCGGACCCCT TCCATCAAGG AGGCGTCCCT CTCAGACAGA 6060  
GGCGTCTATA AGTGCGTGGC CAGCAATGCA GCCGGGGCGG ACAGCCTGGC CATCCGCTG 6120  
CACGTGGCGG CACTGCCCCC CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC 6180  
CCGGGGCTCA GCATTACAT TCACTGCATC CCCAAGGCTG CGCCCTGCTC CAGCGTGC GC 6240  
TGGGTGCTCG GGGACGGTAC TGGATCCGCG CCTCTGCGAGT TCCTCCACGG GAAGTTGTTT 6300  
GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC CCAAGGACAG CGGGCGCTAT 6360  
GAGTGCCTGG CCGCCAACCT GGTAGGCTCC GCAGCGAGGA CCGTGCAGCT GAACGTGCAG 6420  
CGTGCAGCAG CCAACGCGCG CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC 6480  
GGAGGAACCC TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG 6540  
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG AATCAAGGTG 6600  
TTTGCCAATG GGCACCTGGT GGTGAAATCA GTGACGGACA AAGATGCCGG AGATTACCTG 6660  
TGCGTAGCTC GAAATAAGGT TGGTGATGAC TACGTGGTGC TCAAAGTGGG TGTGGTGATG 6720  
AAACCGGCCA AGATTGAACA CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC 6780  
CTGAAAGTGG ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA 6840  
GACGGGAGTC TGGTGAACCT CTTTCATGCAG TCGGATGACA GCGGTGGACG CACCAAGCGC 6900  
TATGTCTGCT TCAACAATGG GACACTCTAC TTAAACGAAG TGGGGATGAG GGAGGAAGGA 6960  
GACTACACCT GCTTTGCTGA AAATCAGGTC GGGAGGACG AGATGAGAGT CAGAGTCAAG 7020  
GTGGTGACAG CGCCCGCCAC ATCCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT 7080  
GGAGACGTGG TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG 7140  
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT ATACCAAGAT 7200  
GGCACTCTCC TTATTCAAGG AGCCCGAGCG TCTGACAGCG GCAACTACAC CTGCCTGGTC 7260  
AGGAACAGCG CCGGAGAGGA TAGGAAGACG GTGTGGATTG ACGTCAACGT CCAGCCACCC 7320  
AAGATCAACG GTAACCCCAA CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT 7380  
CGGAACCTGA TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT 7440  
CCCGAGGGTG TGGTCTGACC AGCTCCATAC TATGGAACC GGATCACTGT CCATGGCAAC 7500  
GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG TCCAGCTGGT ATGCATGGCA 7560  
CGCAACGAGG GAGGGGAGGC GAGGTTGATC GTGCAGCTCA CTGTCTGGA GCCCATGGAG 7620  
AAACCCATCT TCCACGACCC GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC 7680  
AGCCTCAACT GCTCTGCCG GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT 7740  
GGCACCAGTC TGCAGAGTGG ACAGCGCTTCT CAGCGCTTCT ACCACAAGGC TGACGGCATG 7800  
CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT ACCGCTGCGT GGCCCGCAAT 7860  
GCCGCTGGCC ACACGGAGAG GCTGGTCTCC CTGAAGGTGG GACTGAAGCC AGAAGCAAAC 7920  
AAGCAGTATC ATAACCTGGT CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC 7980  
CCTCCCGGGG CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG 8040  
GGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC GGTTCGTGAG 8100  
GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG AGACGGAGTA CGGCCCTTCG 8160  
GTCACCAGCA TCCCCGTGAT TGTGATCGCC TATCCTCCCC GGATCACCAG CGAGCCACC 8220  
CCGGTCATCT ACACCCGGCC CGGGAACACC GTGAACTGA ACTGCATGGC TATGGGGATT 8280  
CCCAAAGCTG ACATCAGGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTTCA 8340  
GCTCGTCTGT ATGGAAACAG ATTTCTTCC CCCCAGGGAT CACTGACCAT CCAGCATGCC 8400  
ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA ACATTCTCGG CAGTGAATCC 8460  
AAAACAACCT ACATCCACGT CTTCTGAAAT GTGGATTCCA GAATGATTGC TTAGGAACTG 8520  
ACAACAAAGC GGGGTTTGTG AGGGAAAGCA GGTGGGGAA TAGGAGCTCT TAAATAATGT 8580  
GTCACAGTGC ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT 8640  
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT GAGACACTTT 8700  
CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG TCTGTGCTCT GACTGCAATT 8760  
TTTCTTCTTT TGCAAATGCC ACTCGACTGC CTTTATAAGC GTCCATAGGA TATCTGAGGA 8820  
ACATTATCA AAAATAAGCC ATAGACCTC ACAACACCTC ACTACCCAT TGAAGACGCA 8880  
TCACCTAGTT AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCAGAT TGACAAGTCA 8940  
TCTTTCAGTT ATTTCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG GATTTAGAAC 9000  
CAGAGTGAAT GATATATATA TATATATTTT AATTCAGAGT TACATACATA CAGCTACCAT 9060  
TTTATATGAA AAAAGAAAAA CATTTCTTCC TGGAACCTCAC TTTTATATA ATGTTTTATA 9120  
TATATATTTT TTCCTTTCAA ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT 9180  
ATTAATAATTA ATAAATTATT GGTCTTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA 9240  
ATATAATTTT AAAAAATTTT TCTCCAAAT CTTTCAAAT CAGTACCAC TGTATATTA 9300  
CCTTCTCCAG GAACCTCCA GTGGGGAAGG CTGCGATATT AGATTTCTT GTATGCAAAG 9360  
TTTTTGTGTA AAGCTGTGCT CAGAGGAGGT GAGAGGAGAG GAAGGAGAAA ACTGCATCAT 9420  
AACTTTACAG AATGAATCT AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT 9480  
CTGGCTTGT CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC 9540  
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT TAAGGTCAAT 9600  
ATACTGTACA TTTGATAATA AAATAATATT CTCCCCAAAA AAAAA

Seq ID NO: 417 Protein sequence  
Protein Accession #: NP\_056234.1

85

1	11	21	31	41	51	
MPKRAHWGAL	SVVLILLWGH	PRVALACPHP	CACYVPSEVH	CTFRSLASVP	AGIARHVERI	60
NLGFNSIQAL	SETSFAGLTK	LELLMIHGNE	IPSIDPGALR	DLSSLQVFKF	SYNKL RVITG	120
QTLQGLSNLM	RLHIDHNKIE	FIHPQAFNGL	TSLRLHLLEG	NLLHQLHPST	FSTFTFLDYF	180
RLSTIRHLYL	AENMVRTLPA	SMLRNMPLLE	NLYLQGNPWT	CDCMRWFLE	WDAKSRGILK	240



	CKKDKAYEGG	QLCAMCFSPK	KLYKHEIHLK	KDMTCLKPSI	ESPLRQNRSR	SIEEEQEQUE	300
	DGGSQILILEK	FQLPQWSISL	NMTDEHGNMV	NLVCDIKKPM	DVYKIHLNQT	DPDIDINAT	360
	VALDFECPMT	RENYEKLWKL	IAYYSEVPVK	LHRELMLSKD	PRVSYQYRQD	ADEEALYYTG	420
5	VRAQILAEPE	WVMQPSIDIQ	LNRRQSTAKK	VLLSYYTQYS	QTISTKDTRQ	ARGRSWVMIE	480
	PSGAVQRDQT	VLEGGPCQLS	CNVKASESPS	IFWVLPDGS	LKAPMDDPDS	KFSILSSGWL	540
	RIKSMEPSDS	GLYQCIQVR	DEMDRMVYRV	LVQSPSTQPA	EKDTVTIGKN	PGESVTLPCN	600
	ALAIPEAHL	WILPNRRIIN	DLANTSHVYM	LPNGTSLIPK	VQVSDSGYYR	CVAVNQGGAD	660
	HFTVGITVTK	KSGSLPSKRG	RRPGAKALSR	VREDIVEDEG	GSGMGDEENT	SRLLHHPKDQ	720
10	EVFLKTKDDA	INGDKKAKKG	RRKLKLWKHS	EKEPETNVAE	GRRVFESRRR	INMANKQINP	780
	ERWADILAKV	RGKNLPKGTE	VPPLIKTTSP	PSLSLEVTPP	FPAVSPPSAS	PVQTVTSAAE	840
	SSADVPLLGE	EEHVLGTISS	ASMGLEHNHN	GVILVEPEVT	STPLEEVVDD	LSEKTEEITS	900
	TEGDLKGTA	PTLISEPYEP	SPTLHTLDTV	YEKPTHEETA	TEGWSAADVG	SSPEPTSSEY	960
	EPPLDAVSLA	ESEPMQYFDP	DLETQSQDE	DKMKEDTFAH	LTPTPTIWN	DSSTSOLFED	1020
	STIGEPGVPG	QSHLQGLTDN	IHLVKSSSLT	QDTLLIKKGM	KEMSQTLOGG	NMLEGDPHNS	1080
15	RSSESEGGES	KSITLPDSTL	GIMSSMSPVK	KPAETTVGTL	LDKDTTIVTT	TPRQKVAPSS	1140
	TMSTHPSRRR	PNGRRRLRPN	KFRHRHKQTP	PTTFAPSETF	STQPTQAPDI	KISSQVESSL	1200
	VPTAWVDNTV	NTPKQLEMEK	NAEPTSKGTP	RRKHGKRPNK	HRYTPSTVSS	RASGSKPSPS	1260
	PENKHRNIVT	PSSETILLPR	TVSLKTEGYP	DSLDMYTTTR	KIYSSYPKVQ	ETLPVTKYKT	1320
	SDGKEIKDDV	ATNVDKHKSD	ILVTGESITN	AIPTSRSLVS	TMGEFKEESS	PVGFPPTPTW	1380
20	NPSRTAQFGR	LQTDIPVTTT	GENLTDPPLL	KELEDVDFTS	EFLSSLTVST	PFHQEEAGSS	1440
	TTLSSIKVEV	ASSQAETTTT	DQDHELTVA	ILLSETRPQN	HTPTAARMKE	PASSSPSTIL	1500
	MSLGQTTTTK	PALPSPRISQ	ASRDSKENVF	LNLYVGNPETE	ATPVNNEGTO	HMSGPNELST	1560
	PSSDRDAFNL	STKLELEKQV	FGSRSLPRGP	DSQRQDGRVH	ASHQLTRVPA	KPILPTATVR	1620
25	LPEMSTQSAS	RYFVTSQSPR	HWTNKPEITT	YPSGALPENK	QFTTPRLSST	TIPLPLHMSK	1680
	PSIPSKFTDR	RTDQFNGYSK	VFGNNNIPEA	RNPVGKPPSP	RIPHYSNGRL	PFFTNKTLSF	1740
	PQLGVTRRPQ	IPTSPAPVMR	ERKVIPIGSYN	RIHSHSTFHL	DFGPPAPPLL	HTPQTGSPS	1800
	TNLQNIPIVS	STQSSISFIT	SSVQSSGSPH	QSSSKFFAGG	PPASKFWSLG	EKPQILTKSP	1860
	QTVSVTAETD	TVFPCEATGK	PKPFVTWTKV	STGALMTFNT	RIQRFEVLKN	GTLVIRKVQV	1920
30	QDRGQYMCAT	SNLHGLDRMV	VLLSVTVQQP	QILASHYQDV	TVYLGDTIAM	ECLAKGTPAP	1980
	QISWIFPDRR	VWQTVSPVES	RITLHENRTL	SIKEASFSDR	GVYKCVASNA	AGADSLAIRL	2040
	HVAALPVIH	QEKLENISLP	PGLSIHIHCT	AKAAPLPSVR	WVLGDGTQIR	PSQFLHGNLF	2100
	VFPNGTLYIR	NLAPKDSGRY	ECVAANLVGS	ARRTVQLNVQ	RAAANARITG	TSPRRTDVRY	2160
	GGTLKLDSCA	SGDPWPRIIL	RLPSKRMIDA	LFSFDSRIKV	FANGTLVVK	VTDKADAGDY	2220
35	CVARNVKGGD	YVVLKVDVVM	KPAKIEHKKE	NDHKVIFYGGD	LKVDCAVATGL	PNPEISWSLP	2280
	DGSLVNSFMQ	SDDSGGRTRK	YVVFNGTLY	FNEVGMREEG	DYTCFAENQV	GKDEMVRVK	2340
	VVTAPATIRN	KTYLAVQVPY	GDVVTVACEA	KGEPMPKVTV	LSPTNKVIPT	SSEKYQIYQD	2400
	GTLTIQKQAR	SDSGNYTCLV	RNSAGEDRKT	VWLVHNVQPP	KINGNPNPIT	TVREIAAGGS	2460
	RKLIDCKAEG	IPTPRVLWAF	PEGVVLPAFY	YGNRITVHGN	GSLDIRSLRK	SDSVQLVCMA	2520
	RNEGGEARLI	VQLTVLEPME	KPIFHDPISE	KITAMAGHTI	SLNCSAAGTP	TPSLVWVLPN	2580
40	GTDLQSGQQL	QRFYHKADGM	LHISGLSSVD	AGAYRCVARN	AAGHTERLVS	LKVGLKPEAN	2640
	KQYHNLVSI	NGETLKLPC	PPGAGQGRFS	WTLPLNGMHLE	GPQTLGRVSL	LDNGTLTVRE	2700
	ASVFDRTGYV	CRMETEYGPS	VTSIPVIVIA	YPPRITSEPT	PVIYTRPGNT	VKLNCMAMGI	2760
	PKADITWELP	DKSHLKAGVQ	ARLYGNRFLH	PQGSLLTIQHA	TQRDAGFYKC	MAKNILGSDS	2820
45	KTTYIHVF						

Seq ID NO: 418 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..5001

50	1	11	21	31	41	51	
	ATGCCAGGCA	CAAACTAAT	CCGAACAGGC	GCCCCAGCAG	ACTACAGAGT	GATATTGAAG	60
	ACCTCTCAAG	AGGACGAATT	GGATGTACCT	GACGACATCA	GCGTCCGGGT	TATGTCATCT	120
55	CAGTCTGTGC	TTGTGTCCTG	GGTGGATCCT	GTTCTGGAAA	AACAGAAGAA	AGTTGTTGCA	180
	TCAAGACAGT	ACACCGTGCG	CTATCGAGAG	AAGGGGGAAT	TGGCCAGGTG	GGATTATAAG	240
	CAGATCGCTA	ACAGGCGTGT	GCTGATTGAG	AACCTGATTC	CAGACACTGT	GTATGAATTT	300
	GCAGTCCGTA	TTTCACAGGG	TGAAAGAGAT	GGCAATGGA	GTACGTCAGT	CTTCCAAAGA	360
	ACACCAGAAT	CTGCCCCTAC	CACAGCTCCT	GAAACTTGA	ACGTCCTGGC	AGTCAATGGC	420
60	AAACCTACAG	TTGTGCTGCT	ATCTTGCGAT	GCGCTACCAG	AGACTGAGGG	GAAAGTGAAA	480
	GTCTGTCTGC	TGGACACAGG	ACTGTTTTCA	GTTTCTCCTC	TCCAACCATC	TGCCAAATCA	540
	TTTCAGAAAT	CATTCTTTCA	TACGCCCCGG	CTCTCAAACC	ATTTGGAGCA	AAGTCCCTCA	600
	CCTATCCTGG	AGACACTACT	TCTGCCCTGG	TGGATGGTCT	GCAGCCTGGG	GAACGCTATC	660
	TTTTCAAAAT	CCGGGCCACA	AACAGGAGAG	GCCTGGGACC	TCACTCCAAA	GCCTTCATTG	720
65	TCGCTATGCC	AACAAGAATG	CAGCTGTACC	CAGAAGGATT	TCAGTTGTCT	AGCTTACCTG	780
	ATCGATATCC	AAACCAAAAC	AGTTAATAAA	GATCCACAAC	TGGAAGGGAG	TGTTTTTGGA	840
	CCATGTTTTC	TTTTCTACTT	CCTCAATTTT	ATGCTGGATA	TTGGCGGCTT	TTCCTTCATT	900
	ATGTGCTATG	AAGACCCANN	TGTTTCTTCT	TTGACAGGCA	ATTCTTTAAA	ATCTGTTGCA	960
	GCCAGTAAGG	CGGATGTTCA	GCAGAACACG	GAGGACAATG	GGAAACCCGA	AAAACCTGAG	1020
70	CCTTCCTCAC	CTTCTCCAG	AGCTCCAGCT	TCCTCCCAAC	ACCCCTCTGT	GCCTGCTTCT	1080
	CCCCAAGGGA	GAAATGCCAA	GGACCTTCTT	CTTGACTTGA	AGAACAAAAT	ATTGGCTAAT	1140
	GGTGGGGCGC	CCCGAAAACC	CCAGCTTCGC	GCCAAAGAGG	CAGAGGAGCT	GGATCTTCAG	1200
	TCGACAGAAA	TCACTGGGGA	TGAGTGGGCT	GGTTCCCGGG	AGGACTCGCC	CATGTCACCC	1260
	TCAGACACCC	AAGACCAGAA	ACGGACCCTG	AGGCCGCCAA	GAGACACGG	CCACTCGGTG	1320
75	GTTGCTCCCG	GCAGGACTGC	AGTGAGGGCC	CGGATGCCAG	CGCTGCCCCG	AAGGGAAGGC	1380
	GTAGATAAGC	CTGGCTTTTC	CCTGGCCACG	CAGCCCCGCC	CAGGGGCGCC	CCCCTCGGCT	1440
	TCGGCTCTC	CTGCCACCA	CGCGTCCACC	CAGGGCACCT	CTCATCGTCC	TTCCCTGCTT	1500
	GCCAGCTTGA	ATGACAACGA	CTTGGTGGAG	TCAGACGAAG	ATGAGCGGCG	TGTGGGCTCC	1560
	CTCCACCCCA	AGGGCGCCTT	CGCCACGCC	CGGTCACGCC	TGTCCCCAG	CCGCCAGTCC	1620
	CCGTCCAGCG	TTCTCCGCGA	CAGAAGCTCT	GTGCACCCCG	GCGCAAAGCC	AGCCTCGCCG	1680
80	GCGCGGAGGA	CCCCCATTC	AGGGGCCGCA	GAGGAAGATT	CCAGTGCCTC	AGCCCCACCC	1740
	TCAAGACTTT	TTCCACCCCA	TGGGGGATCA	TCTCGGCTGC	TGCCCCACCA	GCCACACCTG	1800
	AGCTCTCCAC	TTTCCAAGGG	CGGGAAGGAT	GGTGAGGACG	CCCCAGCCAC	CAACTCCAAT	1860
	GCGCCATCAC	GGTCCACCAT	GTCCTCCTCT	GTCTCTTCTC	ATCTCTCGTC	CAGGACGCAG	1920
85	GTCTCTGAGG	GAGCGGAGGC	TTCTCTGAGT	GAAAGCCACG	GTGACGGCGA	TAGGGAAGAC	1980
	GGCGGAAGGC	AGGCGGAGGC	CACGGCCACG	ACGCTGCGGG	CCCGGCTGCT	CTCTGGACAC	2040
	TTCCATTTGC	TCAGACACAA	ACCCTTTGCT	GCCAACGGGA	GGTCTCCAAG	CAGGTTTCAGC	2100
	ATTGGGCGGG	GACCTCGGCT	GCAGCCCTCC	AGCTCCCCAC	AGTCGACTGT	GCCCTCCCGA	2160



	GCCCACCCCA	GGGTTCCCTC	TCACTCTGAT	TCCCACCCCTA	AGCTTAGCTC	AGGTATCCAT	2220
	GGAGACGAGG	AGGATGAGAA	GCCGCTTCCT	GCCACCGTTG	TCAATGACCA	CGTGCCTTCC	2280
	TCCTCCAGGC	AGCCCATCTC	CCGGGGCTGG	GAGGACTTAA	GGAGAAGCCC	GCAGAGAGGG	2340
5	GCCAGCCTGC	ATCGGAAGGA	ACCCATCCCA	GAGAACCCCA	AATCCACAGG	GGCAGATACA	2400
	CATCCTCAGG	GCAAGTACTC	CTCCCTGGCC	TCCAAGGCTC	AGGATGTTCA	ACAGAGCACA	2460
	GACGCGGACA	CGGAGGGTCA	TTCTCCCAAA	GCACAGCCAG	GGTCCACAGA	CCGCCACGCG	2520
	TCCCTTGCTC	GTCTCCCGC	AGCACGGTCA	CAGCAGCATC	CCAGTGTTCC	CAGAAGGATG	2580
	ACACCCGGCC	GGGCCCCAGA	ACAGCAGCCC	CCTCCTCCCG	TCGCCACGTC	CCAGCACCAC	2640
10	CCGGGACCCC	AGAGCAGAGA	CGCGGGTCGG	TCACCTTCCC	AGCCCAGGCT	CTCACTGACC	2700
	CAGGCCGGGC	GGCCCCGCCC	CACGTCGCAG	GGCCGCTCCC	ACTCCTCCTC	GGACCCCTAC	2760
	ACGGCGAGCT	CCAGAGGGAT	GCTCCCCACG	GCCCTCCAGA	ACCAGGACGA	GGATGCCCAG	2820
	GGCAGCTACG	ACGACGACAG	CACAGAAATC	GAGGCCCAGG	ATGTGCGGGC	CCCCGCGCAC	2880
	GCCGCGCGCG	CCAAGGAGGC	AGCTGCGTCC	CTTCCCAAGC	ACCAGCAGGT	GGAGTCTCCC	2940
15	ACAGGCGCAG	GGGCAGGTGG	CGACCACAGG	TCCCAGCGCG	GACATGCGGC	CTCCCCCGCC	3000
	AGGCCAGGCC	GACCCGCGCG	CCCCCAGTCC	CGCGCCCGGG	TCCCCAGCAG	GGCAGCGCCG	3060
	GGGAAGTCGG	AGCCTCCTTC	CAAGCGGCCC	CTGTCTCCA	AGTCCCAGCA	GTGCGTCTCA	3120
	GCCGAGGACG	AGGAGGAGGA	GGACGCGGGG	TTTTTTAAAG	GCGGGAAAGA	AGACCTTCTG	3180
	TCTTCTCTG	TGCCAAAGTG	GCCCTCTTCC	TCCACTCCCA	GGGGCGGCAA	AGACGCCGAT	3240
20	GGGAGCCTCG	CCAAGGAAGA	GAGGGAGCCT	GCCATCGCGC	TTGCCCCCTG	CGGAGGGAGC	3300
	CTGGCTCCTG	TGAAGCGACC	TCTCCCCCAG	CCTCCAGGCA	GCTCCCCCAG	GGCCTCCAC	3360
	GTCCCTTCCC	GACCGCCGCC	TGCAGCGCT	GCCACCGTGA	GCCCGTCCG	GGGCACCCAC	3420
	CCCTGGCCGC	GGTACACCAC	GCGCGCCCCV	CCTGGCCACT	TCTCCACCAC	CCCGATGCTG	3480
	TCCTTGCGCC	AGAGGATGAT	GCATGCCAGA	TTCCGTAACC	CTCTCTCCCG	ACAGCCTGCC	3540
25	AGACCTCTT	ACAGACAAGG	TTATAATGGC	AGACCAAATG	TAGAAGGGAA	AGTCCTTCCT	3600
	GGTAGTAATG	GAAAAACGAA	TGGACAGAGA	ATTATCAATG	GCCCTCAAGG	AACAAAGTGG	3660
	GTTGTGGACC	TTGATCGTGG	GTTAGTATTG	AATGCAGAAG	GAAGGTACCT	CCAAGATTCA	3720
	CATGGAAATC	CTCTTCGGAT	TAAACTAGGA	GGAGATGGTC	GAACCATTTG	AGATCTGGAA	3780
	GGGACCCCGC	TGGTGAGTCC	TGACGGCCTC	CCACTCTTTG	GGCAGGGGCG	ACATGGCACA	3840
30	CCTCTGGCCA	ATGCCCAAGA	TAAGCCAATT	TTGAGTCTTG	GAGGAAAGCC	GCTGGTGGGC	3900
	TTGGAGGTCA	TCAAAAAAAC	CACCCATCCC	CCTACCACTA	CCATGCAGCC	CACCACTACT	3960
	ACGACGCCCC	TGCCTACCAC	TACAACCCCG	AGGCCACCA	CTGCCACCAC	CATGCAGCCC	4020
	ACCACTACTA	CGACGCCCCCT	GCCTACCACT	ACACCGAGGC	CCACCACTGC	CACCACCCGC	4080
	CGCACGACCA	CCAGGCGTCC	AACAACCA	GTCCGAACCA	CTACGCGGAC	AACCACCACC	4140
35	ACCACCCCA	AACCCACCAC	TCCCCATCCC	ACCTGTCCCC	CTGGGACCTT	GGAACGGCAC	4200
	GACGATGATG	GCAACCTGAT	AATGAGCTCC	AATGGGATCC	CAGAGTGCTA	CGCTGAAGAA	4260
	GATGAGTTCT	CAGGCTTGGA	GACTGACACT	GCAGTACCTA	CGGAAGAGGC	CTACGTTATA	4320
	TATGATGAAG	ATTATGAATT	TGAGACGTCA	AGGCCACCAA	CCACCACTGA	GCCTTCGACC	4380
	ACTGCTACCA	CACCGAGGGT	GATCCCAGAG	GAAGGCGCCA	TCAGTTCCTT	TCCTGAAGAA	4440
40	GAATTTGATC	TGGCTGGAAG	GAAACGATTT	GTTGCTCCTT	ACGTGACGTA	CCTAAATAAA	4500
	GACCCATCAG	CCCCGTGCTC	TCTGACTGAT	GCCACTGGATC	ACTTCCAAGT	GGACAGCCTG	4560
	GATGAAATCA	TCCCCAATGA	CCTGAAGAAG	AGTGATCTGC	CTCCCCAGCA	TGCTCCCCGC	4620
	AACATCACCG	TGGTGGCCGT	GGAAGGTTGC	CACTCATTTG	TCATTGTGGA	TTGGGACAAA	4680
	GCCACCCAG	GAGATTTGGT	CACAGGTTAT	TTGGTTTACA	GTGCATCCTA	TGAAGATTTT	4740
45	ATCAGGAACA	AGTTTTCCAC	TCAAGCTTCA	TCAGTAACTC	ACTTGCCCAT	TGAGAACCCTA	4800
	AAGCCCAACA	CGAGGTATTA	TTTTAAAGTG	CAAGCACAAA	ATCCTCATGG	CTACGGACCT	4860
	ATCAGCCCTT	CGGTCTCATT	TGTACCGGAA	TCAGATAATC	CTCTGCTTGT	TGTGAGGCC	4920
	CCAGGCGGTG	AGCTATCTGG	ATCCCATTCG	CTTTCAAACA	TGATCCCAGC	TACACGGACT	4980
	GCCATGGACG	GCAATATGTG	AAGCGCACGT	GGTATCGAAA	GTTCTGTGGA	GTTGTTCTTT	5040
50	GTAATTCAT	GAGGTATAAA	ATCTACCTCA	GTGACAACTT	GAAAGATACA	TTCTACAGCA	5100
	TTGGAGACAG	CTGGGGAAGA	GGTGAAGACC	ATTGCCAATT	TGTGGATTCA	CACCTTGATG	5160
	GAAGAACAGG	GCCTCAGTCC	TATGTAGAAG	CCCTCCCTAC	TATTCAAGGC	TACTATCGCC	5220
	AGTATCGTCA	GGAGCCTGTC	AGGTTTGGGA	ACATCGGCTT	CGGAACCCCT	TACTACTATG	5280
	TGGGCTGTA	CGAGTGTGGG	GTCTCCATCC	CTGGAAAGTG	GTAATCACAG	GACCGTCATG	5340
55	CTGCAAGCTT	GCCCTGCCCA	GCCCCACCAA	CTAAGTCGCA	CTAGGGGCTG	TGAGCAAAGA	5400
	CAGCCAGCAT	GCTCAGCCCC	GCTGCCCTAG	GTGCCAGGAA	GGTCACAGAT	GGACACTGGC	5460
	CATTCTGGTC	ATCTCAGTCT	GGAACCTCAGT	CCCCTCTCTT	GGCCTGGACA	ATGAACAGGA	5520
	TTCACTTTTG	CTGTAACTT	TGCTTCTCTA	CTTTTTTTTG	TTTGTGTTGA	ATAGCACATC	5580
	CCAGAGACAT	CAGAAACCAG	CAACTGATTC	AGTGATGATT	CCCAGACTTT	TTAGGCATGA	5640
60	AATTCGGACA	CTTCAGTATT	TCCAGGAATA	GCATATGCAC	GCTGTTCTTG	CTTCATGGAA	5700
	TGCTACATGC	TTTCTGTTTT	TCTCATTTTG	GATTTCTCCA	AAACTAACTG	AATTTAAGCT	5760
	TCAGGTCCCT	TTGTATGCAG	TAGAAAGGAA	TTATTAATAA	CACCACCAA	GAAAAATAAT	5820
	ATATCCTACT	TGAAATTTAC	TCTATGGACT	TACCCACTGC	TAGAATAAAT	GTATCAAATC	5880
	TTATTTGTAA	ATTCTCAATT	TTGATATATA	TATGTATATA	TGCATATACA	TATCCACACT	5940
65	TGCTGCAAG	AATATTGATT	AAAATTGCTA	AATTTGTACT	TGTTACCAA	AAAAAAAAAA	6000
	AAAAAA						

Seq ID NO: 419 Protein sequence  
Protein Accession #: Eos sequence

70	1	11	21	31	41	51	
	MPGTLKLR	APADYRVILK	TSQEDELDP	DDISVRVMSS	QSVLVSWVDP	VLEKQKKVVA	60
	SRQYTVRYRE	KGELARWDYK	QIANRRVLIE	NLIPDTVYEF	AVRISQGERD	GKWSTSVFQR	120
75	TPESAPTTAP	ENLNVWPVNG	KPTVVAASWD	ALPETEGVK	VCLLDITGLFS	VSSFQPSAKS	180
	FQNTFFHTPR	LSNHLEQSPS	PILETLLLPW	WMVCSLGNAI	FSKSGPQTGE	AWDLTPKPSL	240
	SLCQQECSCT	QKDFSCLAYL	IDIQTKQVNK	DPQLEGSVFG	PCFLFYFLTF	MLDIGGFSFI	300
	MCYEDPVSSL	TGNSLKSVA	SKADVQNTRE	DNGKPEKPEP	SSPSRAPAS	SQHPSVPASP	360
	QGRNAKDLLL	DLKNKILANG	GAPRKPLRA	KKAEELDLQS	TEITGEEELG	SREDSPMSPS	420
80	DTQDQKRTL	PPSRHGHSVV	APGRTAVRAR	MPALPRREGV	DKPGFSLATQ	PRPGAPPSAS	480
	ASPAHASTQ	GTSHRPSLPA	SLNDNDLVDS	DEDERAVGSL	HPKGAFAPQR	PALSPSRQSP	540
	SSVLRDRSSV	HPGAKPASPA	RRTPHSGAAE	EDSSASAPPS	RLSPPHGGSS	RLLPTQPHLS	600
	SPLSKGKDG	EDAPATNSNA	PSRSTMSSSV	SSHLSSRTQV	SEGAEASDGE	SHGDGDREDG	660
	GRQAEATAQT	LRARPASGHF	HLLRHKPFPA	NGRSPSRFSI	GRGPRLQPSS	SPQSTVPSRA	720
85	HPRVPSHSDS	HPKLSSGIHG	DEEDEKPLPA	TVVNDHVPSS	SRQPISRGE	DLRRSPQRGA	780
	SLHRKEPIPE	NPKSTGADTH	PQGYSSLAS	KAQDVQSTD	ADTEGHSPKA	QPGSTDRHAS	840
	PARPPAARSQ	QHPSVPRRMT	PGRAPQPP	PPVATSQHHP	GPQSRDAGRS	PSQPRLSLTQ	900
	AGRPRPTSQG	RSHSSSDPYT	ASSRGMPLTA	LQNQDEDAQG	SYDDDSTEVE	AQDVRAPAH	960



ARAKEAAASL PKHQQVESPT GAGAGGDHRS QRGHAASPAR PSRPGGPQSR ARVPSRAAPG 1020  
 KSEPPSKRPL SSKSQSVSA EDEEEEDAGF FKGGKEDLLS SSVPKWPSSS TPRGGKDADG 1080  
 SLAKEEREPA IALAPRGGS LALAPRGGS APVKRPLPPP PGSSPRASHV PSRPPERSAA TVSPVAGTHP 1140  
 WPRYTTRAPP GHFSTTFMLS LRQRMHARF RNPLSRQPAR PSYRQGYNGR PNVEGKVLPG 1200  
 SNGKPNGQRI INGPQGTKWV VDLDRGLVLN AEGRYLQDSH GNPLRIKLGG DGRTIVDLEG 1260  
 TPVVS PDGLP LFGQGRHGTP LANAQDKPIL SLGGKPLVGL EVIKKTTHPP TTTMQPTTTT 1320  
 TPLPTTTTPR PTTATTMQPT TTTTLPPTTT PRPTTATTRR TTTRRPTTTV RTTTRTTTTT 1380  
 TPKPTTPIPT CPPGTLEHHD DDGNLIMSSN GIPECYAEED EFSGLETDTA VPTEEAYVIY 1440  
 DEDYEFETSR PPTTTEPSTT ATTPRVIPEE GAISFPPEEE FDLAGRKRFEV APYVTYLNKD 1500  
 PSAPCSLTDA LDHFQVDSL D EIIPNDLKKS DLPPQHAPRN ITVVAVEGCH SFVIVDWDKA 1560  
 TPGDLVTGYL VYSASYEDFI RNKFSTQASS VTHLPIENLK PNTRYFVKVQ AQNPHGYGPI 1620  
 SPSVSFVTE DNPLLVVRPP GGELSGSHSL SNMIPATRTA MDGNM

Seq ID NO: 420 DNA sequence  
 Nucleic Acid Accession #: NM\_022743  
 Coding sequence: 128..1237

1 11 21 31 41 51  
 | | | | |  
 GTGGATTTTA GAGATACCTC CCTCCTTCT GCTCAGCTGC CTTGCAGTAA TTAAACTCTT 60  
 TCTCTGCTGC AACACCCCTA CTGTTCTCCG TGTATTGGCT TTTCTGGGCA GCAGGAAGGA 120  
 AAAGCTGATG CGATGCTCTC AGTGCCCGT CGCCAAATAC TGATGTGCTA AGTGTGAGAA 180  
 AAAAGCTTGG CCAGACCACA AGCGGGAATG CAAATGCCTT AAAAGCTGCA AACCAGATA 240  
 TCCTCCAGAC TCCGTTGAC TTCTTGGCAG AGTTGTCTTC AAATTATGG ATGGAGCACC 300  
 TTCAGAATCA GAGAAGCTTT ACTCATTTTA TGATCTGGAG TCAATATTA ACAAACTGAC 360  
 TGAAGATAAG AAAGAGGGCC TCAGGCAACT CGTAATGACA TTTCAACATT TCATGAGAGA 420  
 AGAAATACAG GATGCCTCTC AGCTGCCACC TGCCTTTGAC CTTTTGAAG CCTTTGCAA 480  
 AGTGATCTGC AACTCTTCA CCATCTGTAA TCGGAGATG CAGGAGTTG GTGTTGGCCT 540  
 ATATCCCAGT ATCTCTTTC TCAATCACAG CTGTGACCCC AACTGTTCGA TTGTGTTCAA 600  
 TGGGCCCCAC CTCTTACTGC GAGCAGTCCG AGACATCGAG GTGGGAGAGG AGCTCACCAT 660  
 CTGCTACCTG GATATGCTGA TGACCAGTGA GGAGCGCCGG AAGCAGCTGA GGGACCAGTA 720  
 CTGCTTTGAA TGTGACTGTT TCCGTTGCCA AACCAGGAC AAGGATGCTG ATATGCTAAC 780  
 TGGTGATGAG CAAGTATGGA AGGAAGTTCA AGAATCCCTG AAAAAAATTG AAGAACTGAA 840  
 GGCACACTGG AAGTGGGAGC AGTTCTGGC CATGTGCCAG GCGATCATA GCAGCAATTC 900  
 TGAACGGCTT CCGGATATCA ACATCTACCA GCTGAAGGTG CTCGACTGCG CCATGGATGC 960  
 CTGCATCAAC CTCGGCCTGT TGGAGGAAGC CTTGTTCTAT GGTACTCGGA CCATGGAGCC 1020  
 ATACAGGATT TTTTCCCAG GAAGCCATCC CGTCAGAGGG GTTCAAGTGA TGAAAGTTGG 1080  
 CAAACTGCAG CTACATCAAG GCATGTTTCC CCAAGCAATG AAGAACTGTA GACTGGCTTT 1140  
 TGATATTATG AGAGTGACAC ATGGCAGAGA ACACAGCCTG ATTGAAGATT TGATTCTACT 1200  
 TTTAGAAGAA TGGCAGGCCA ACATCAGAGC ATCCTAAGGG AACGCAGTCA GAGGGAATA 1260  
 CGGCGTGTGT CTTTGTGAA TGCTTATTG AGGTACACA CTCTATGCTT TGTAGCTGT 1320  
 GTGAACCTCT CTTATTGGAA ATTCTGTTCC GTGTTTGTGT AGGTAAATA AGGCAGACAT 1380  
 GGTTTGCAA CCACAAGAA CATTAGTTGT AGAGAAGCAC GATTATAATA AATTCAAAAC 1440  
 ATTTGGTTGA GGATGCCAAA AAAAAAAAAA AAAAAA

Seq ID NO: 421 Protein sequence  
 Protein Accession #: NP\_073580

1 11 21 31 41 51  
 | | | | |  
 MRCSQCRVAK YCSAKCQKKA WPDHKRECKC LKCKPRYP DSVRLGRV FKLMDGAPSE 60  
 SEKLYSFYDL ESNINKLTED KKEGLRQLVM TFQHFMRREE QDASQLPPAF DLFEAFKVI 120  
 CNSLFTNAE MQEVGVGLYP SISLLNHSCD PNCISIVFNGP HLLLRVRDI EVGEELTICY 180  
 LDMLMTSEER RKQLRDQYCF EDCFCRCQTQ DKDADMLTGD EQVWKEVQES LKKIEELKAH 240  
 WKWEQVLAMC QAISSNSER LPDINIYQLK VLDCAMDACI NLGLLEEALF YGTRTMEPYR 300  
 IFFPGSHVVR GVQVMKVGKL QLHQGMFPQA MKNLRLAFDI MRVTHGREHS LIEDLILLLE 360  
 ECDANIRAS

Seq ID NO: 422 DNA sequence  
 Nucleic Acid Accession #: NM\_003014.2  
 Coding sequence: 238..648

1 11 21 31 41 51  
 | | | | |  
 GGCGGGTTTC CGCCCCGAAG GCTGAGAGCT GCGCTGCTC GTGCCCTGTG TGCCAGACGG 60  
 CGGAGCTCCG CGGCGGACC CCGCGGCCCC GCTTTGCTGC CGACTGGAGT TTGGGGGAAG 120  
 AAATCTCTCT GCGCCCCAGA AGATTTCTTC CTCGGCGAAG GGACAGCGAA AGATGAGGGT 180  
 GGCAGGAAGA GAAGGCGCTT TCTGTCTGCC GGGGTGCGAG CGCGAGAGGG CAGTGCCATG 240  
 TTCCTCTCCA TCCTAGTGGC GCTGTGCTTG TGGCTGCACC TGGCGCTGGG CGTGCGCGGC 300  
 GCGCCCTGCG AGGCGGTGCG CATCCCTATG TGCCGGCACA TGCCCTGGAA CATCACGCGG 360  
 ATGCCCAACC ACCTGCACCA CAGCACGCGA GAGAAGCCCA TCCTGGCCAT CGAGCAGTAC 420  
 GAGGAGCTGG TGGACGTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTTCTG TGCCATGTAC 480  
 GCGCCCATTT GCACCTGGA GTTCTGTCAC GACCTATCA AGCCGTGCAA GTCGGTGTGC 540  
 CAACGCGCGC GCGACGACTG CGAGCCCCCT ATGAAGATGT ACAACCACAG CTGCCCCGAA 600  
 AGCCTGGCCT GCGACGAGCT GCCTGTCTAT GACCGTGGCG TGTGCATTTT GCCTGAAGCC 660  
 ATCGTCAACG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACCAGA CATGATGGTA 720  
 CAGGAAAGGC CTCTTGATGT TGAATGTTAA CGCCTAAGCC CCGATCGGTG CAAGTGTA 780  
 AAGGTGAAGC CAATTTGGC AACGATCTCT AGCAAAAAC ACAGCTATGT TATTCATGCC 840  
 AAAATAAAAG CTGTGCAGAG GAGTGGCTGC AATGAGGTCA CAACGGTGGT GGATGTAAAA 900  
 GAGATCTTCA AGTCCTCAT ACCCATCCCT CGAACTCAAG TCCCGCTCAT TACAAATTCT 960  
 TCTTGCCAGT GTCCACACAT CCTGCCCAT CAAGATGTTT TCATCATGTG TTACGAGTGG 1020  
 CGTTCAAGGA TGATGCTTCT TGAAAATTGC TTAGTTGAAA AATGGAGAGA TCAGCTTAGT 1080  
 AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAAG 1140  
 AAGAAAACAG CCGGGCGCAC CAGTCGTAGT AATCCCCCA AACCAAGGG AAAGCCTCCT 1200  
 GCTCCCAAAC CAGCAGTCC CAAGAAGAAC ATTAATACTA GGAGTGCCCA GAAGAGAACA 1260  
 AACCCGAAAA GAGTGTGAGC TAAGTAGTTT CCAAGCGGA GACTTCCGAC TTCCTTACAG 1320  
 GATGAGGCTG GGCATTGCCT GGGACAGCCT ATGTAAGGCC ATGTGCCCT TGCCCTAACA 1380



	ACTCACTGCA	GTGCTCTTCA	TAGACACATC	TTGCAGCATT	TTTCTTAAGG	CTATGCTTCA	1440
	GTTTTTCTTT	GTAAGCCATC	ACAAGCCATA	GTGGTAGGTT	TGCCCCTTGG	TACAGAAGGT	1500
	GAGTTAAAGC	TGGTGGAAAA	GGCTTATTGC	ATTGCATTCA	GAGTAACCTG	TGTGCATACT	1560
5	CTAGAAGAGT	AGGGAAAATA	ATGCTTGTTA	CAATTCGACC	TAATATGTGC	ATTGTAAAAT	1620
	AAATGCCATA	TTTCAAACAA	AACACGTAAT	TTTTTTACAG	TATGTTTTAT	TACCTTTTGA	1680
	TATCTGTTGT	TGCAATGTTA	GTGATGTTTT	AAAATGTGAT	GAAAATATAA	TGTTTTTAAG	1740
	AAGGAACAGT	AGTGAATGA	ATGTTAAAAG	ATCTTTATGT	GTTTATGGTC	TGCAGAAGGA	1800
	TTTTTGATGAT	GAAAGGGGAT	TTTTTGAAAA	ATTAGAGAAG	TAGCATATGG	AAAATTATAA	1860
10	TGTGTTTTTT	TACCAATGAC	TTTCAGTTTCT	GTTTTTAGCT	AGAAACTTAA	AAACAAAAAT	1920
	AATAATAAAG	AAAAATAAAT	AAAAAGGAGA	GGCAGACAAT	GTCTGGATTG	CTGTTTTTTG	1980
	GTTACCTGAT	TTCCATGATC	ATGATGCTTC	TTGTCAACAC	CCTCTTAAGC	AGCACCAGAA	2040
	ACAGTGAGTT	TGTCTGTACC	ATTAGGAGTT	AGGTACTAAT	TAGTTGGCTA	ATGCTCAAGT	2100
	ATTTTATACC	CACAAGAGAG	GTATGTCACT	CATCTTACTT	CCCAGGACAT	CCACCCTGAG	2160
	AATAATTTGA	CAAGCTTAAA	AATGGCCTTC	ATGTGAGTGC	CAAATTTTGT	TTTTCTTCAT	2220
15	TTAAATATTT	TCCTTGCCCTA	AATACATGTG	AGAGGAGTTA	AATATAAATG	TACAGAGAGG	2280
	AAAGTTGAGT	TCCACCTCTG	AAATGAGAAT	TACTTGACAG	TTGGGATACT	TTAATCAGAA	2340
	AAAAAGAACT	TATTTGCAGC	ATTTTATCAA	CAAATTTTCT	AATTGTGGAC	AATTGGAGGC	2400
	ATTTATTTTA	AAAAACAATT	TTATTGGCCT	TTTGCTAACA	CAGTAAGCAT	GTATTTTATA	2460
20	AGGCATTCAA	TAAATGCACA	ACGCCCCAAG	GAAATAAAAT	CCTATCTAAT	CCTACTCTCC	2520
	ACTACACAGA	GGTAATCACT	ATTAGTATTT	TGGCATATTA	TTCTCCAGGT	GTTTGCTTAT	2580
	GCACTTATAA	AATGATTTGA	ACAAATAAAA	CTAGGAACCT	GTATACATGT	GTTTCATAAC	2640
	CTGCCCTCCT	TGCTTGGCCC	TTTATTGAGA	TAAGTTTTCC	TGTCAAGAAA	GCAGAAACCA	2700
	TCTCATTTCT	AACAGCTGTG	TTATATTCCA	TAGTATGCAT	TACTCAACAA	ACTGTTGTGC	2760
25	TATTGGATAC	TTAGGTGGTT	TCTTCACTGA	CAATACTGAA	TAAACATCTC	ACCGGAATTC	

Seq ID NO: 423 Protein sequence  
Protein Accession #: NP\_003005.1

30	1	11	21	31	41	51	
	MFLSILVALC	LWLHLALGVR	GAPCEAVRIP	MCRHMPWNIT	RMPNHLHHST	QENAILAIEQ	60
	YEELVDVNC	AVLRFFFCAM	YAPICTLEFL	HDPIKPCKSV	CQRARDDCEP	LMKMYNHSWP	120
	ESLACDELFP	YDRGVCISPE	AIIVTDLPELV	KWIDITPDM	VQERPLDVDC	KRLSPDRCKC	180
35	KKVKPTLATY	LSKNYSYVIH	AKIKAVQRS	CNEVTTVVDV	KEIFKSSSPI	PRTQVPLITN	240
	SSCQCPLILP	HQDVLIMCYE	WRSRMMLLEN	CLVEKWRDQL	SKRSIQWEER	LQEQRRTVQD	300
	KKKTAGRTSR	SNPPKPKGKP	PAPKPASPKK	NIKTRSAQKR	TNPKRV		

Seq ID NO: 424 DNA sequence  
Nucleic Acid Accession #: BC010423  
Coding sequence: 248..1780

	1	11	21	31	41	51	
45	CACAGCGTGG	GAAGCAGCTC	TGGGGGAGCT	CGGAGCTCCC	GATCACGGCT	TCTTGGGGGT	60
	AGCTACGGCT	GGGTGTGTAG	AACGGGGCCG	GGGCTGGGGC	TGGGTCCCTT	AGTGGAGACC	120
	CAAGTGCAG	AGGCAAGAAC	TCTGCAGCTT	CCTGCCTTCT	GGGTCACTTC	CTTATTCAAG	180
	TCTGCAGCCG	GCTCCCAGGG	AGATCTCGGT	GGAACCTCAG	AAACGCTGGG	CAGTCTGCCT	240
	TTCAACCATG	CCCCTGTCCC	TGGGAGCCGA	GATGTGGGGG	CCTGAGGCCT	GGCTGCTGCT	300
50	GCTGCTACTG	CTGGCATCAT	TTACAGGCCG	GTGCCCCGCG	GGTGAGCTGG	AGACCTCAGA	360
	CGTGGTAAC	GTGGTGCTGG	GCCAGGACGC	AAAACCTGCC	TGCTTCTACC	GAGGGGACTC	420
	CGGCGAGCAA	GTGGGGCAAG	TGGCATGGGC	TCGGGTGGAC	GCGGGCGAAG	GCGCCCAGGA	480
	ACTAGCGCTA	CTGCACTCCA	AATACGGGCT	TCATGTGAGC	CCGGCTTACG	AGGGCCGCGT	540
	GGAGCAGCCG	CGCCCCCCAC	GCAACCCCTT	GGACGGCTCA	GTGCTCCTGC	GCAACGCAGT	600
55	GCAGGCGGAT	GAGGGCGAGT	ACGAGTGCCG	GGTCAGCACC	TTCCCCGCCG	GCAGCTTCCA	660
	GGCGCGGCTG	CGGCTCCGAG	TGCTGGTGCC	TCCCCTGCCC	TCACTGAATC	CTGGTCCAGC	720
	ACTAGAAGAG	GGCCAGGGCC	TGACCTGGC	AGCCTCCTGC	ACAGCTGAGG	GCAGCCCAGC	780
	CCCCAGCGTG	ACCTGGGACA	CGGAGGTCAA	AGGCACAACG	TCCAGCCGTT	CCTTCAAGCA	840
	CTCCCGCTCT	GCTGCCGTCA	CCTCAGAGTT	CCACTTGGTG	CCTAGCCGCA	GCATGAATGG	900
60	GCAGCCACTG	ACTTGTGTGG	TGTCCCCTCC	TGGCCTGCTC	CAGGACCAAA	GGATCACCCA	960
	CATCCTCCAC	GTGTCCTTCC	TTGCTGAGGC	CTCTGTGAGG	GGCCTTGAAG	ACCAAAATCT	1020
	GTGGCACATT	GGCAGAGAAG	GAGCTATGCT	CAAGTGCCCTG	AGTGAAGGGC	AGCCCCCTCC	1080
	CTCATACAAC	TGGACACGGC	TGGATGGGCC	TCTGCCCAGT	GGGGTACGAG	TGGATGGGGA	1140
	CACTTTGGGC	TTTCCCCCAC	TGACCACTGA	GCACAGCGGC	ATCTACGTCT	GCCATGTCAG	1200
65	CAATGAGTTC	TCCTCAAGGG	ATTCTCAGGT	CACTGTGGAT	GTTCTTGACC	CCCAGGAAGA	1260
	CTCTGGGAAG	CAGGTGGACC	CTCGGTGGTG	CTCGGTGGTG	GTGGTGGGTG	TGATCGCCGC	1320
	ACTCTTGTTT	TGCTTCTTGG	TGGTGGTGGT	GGTGCTCATG	TCCCGATACC	ATCGGCGCAA	1380
	GGCCCAGCAG	ATGACCCAGA	AATATGAGGA	GGAGCTGACC	CTGACCAGGG	AGAACTCCAT	1440
	CCGGAGGCTG	CATTCCCATC	ACACGGACCC	CAGGAGCCAG	CCGGAGGAGA	GTGTAGGGCT	1500
70	GAGAGCCGAG	GGCCACCCTG	ATAGTCTCAA	GGACAACAGT	AGCTGCTCTG	TGATGAGTGA	1560
	AGAGCCCAG	GGCCGCAAGT	ACTCCACGCT	GACCACGGTG	AGGGAGATAG	AAACACAGAC	1620*
	TGAAGTGGCT	TCTCCAGGCT	CTGGGCGGGC	CGAGGAGGAG	GAAGATCAGG	ATGAAGGCAT	1680
	CAAAACAGGC	ATGAACCAT	TTGTTTCAGG	GAATGGGACC	CTACGGGCCA	AGCCACGGG	1740
	CAATGGCATC	TACATCAATG	GGCGGGGACA	CCTGGTCTGA	CCCAGGCCTG	CCTCCCTTCC	1800
75	CTAGGCCTGG	CTCCTTCTGT	TGACATGGGA	GATTTTAGCT	CATCTTGGGG	GCCTCCTTAA	1860
	ACACCCCCAT	TTCTTGCGGA	AGATGCTCCC	CATCCCCTG	ACTGCTTGAC	CTTTACCTCC	1920
	AACCCTTCTG	TTCATCGGGA	GGGCTCCACC	AATTGAGTCT	CTCCCACCAT	GCATGCAGGT	1980
	CACGTGTGTG	GTGCATGTGT	GCCTGTGTGA	GTGTTGACTG	ACTGTGTGTG	TGTGGAGGGG	2040
80	TGACTGTCCG	TGGAGGGGTG	ACTGTGTCCG	TGGTGTGTAT	TATGCTGTCA	TATCAGAGTC	2100
	AAGTGAAGT	TGGTGTATGT	GCCACGGGAT	TTGAGTGGTT	GCGTGGGCAA	CACGTGTCAGG	2160
	GTTTGGCGTG	TGTGTATGTG	GGCTGTGTGT	GACCTCTGCC	TGAAAAAGCA	GGTATTTTCT	2220
	CAGACCCAG	AGCAGTATTA	ATGATGCAGA	GGTTGGAGGA	GAGAGGTGGA	GACTGTGGCT	2280
	CAGACCCAGG	TGTGCGGGCA	TAGCTGGAGC	TGGAATCTGC	CTCCGGTGTG	AGGGAACCTG	2340
	TCTCCTACCA	CTTCGGAGCC	ATGGGGGCAA	GTGTGAAGCA	GCCAGTCCCT	GGGTGAGCCA	2400
85	GAGGCTTGAA	CTGTTACAGA	AGCCTCTGCG	CCTCTGGTGG	CCTCTGGGCC	TGCTGCATGT	2460
	ACATATTTTC	TGTAATAATA	CATGCGCCCG	GAGCTTCTTG	CAGGAATACT	GCTCCGAATC	2520
	ACTTTTAATT	TTTTTCTTTT	TTTTTCTTTG	CCCTTTCCAT	TAGTTGTATT	TTTTATTTAT	2580
	TTTTATTTTT	ATTTTTTTTT	AGAGTTTGAG	TCCAGCCTGG	ACGATATAGC	CAGACCCTGT	2640



CTGTAAAAAA ACCAAAACCC AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 425 Protein sequence  
Protein Accession #: AAH10423

5  
10  
15

1	11	21	31	41	51	
MPLSLGAEMW	GPEAWLLLLL	LLASFTGRCP	AGELETSDEV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAWARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRVS	TFFAGSFQAR	LRLRVLPPL	PSLNPGPALE	EGQGLTLAAS	CTAEGSPAPS	180
VTWDTEVKGT	TSSRSFKHSR	SAAVTSEFHL	VPSRSMNGQP	LTCVVSHPLG	LQDQRITHIL	240
HVSFLAEASV	RGLEDQNLWH	IGREGAMKLC	LSEGQPPPSY	NWTRLDGPLP	SGVRVDGDTL	300
GFPPLTTEHS	GIYVCHVSNE	FSSRDSQVTV	DVLDPQEDSG	KQVDLVSASV	VVGVGVIAALL	360
FCLLVVVVVL	MSRYHRRKAQ	QMTQKYEEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
EGHPDSLKDN	SSCSVMSEEP	EGRSYSTLTT	VREIETQTEL	LSFGSGRAEE	EEDQDEGIKQ	480
AMNHVQENG	TLRAKPTGNG	IYINGRGLV				

Seq ID NO: 426 DNA sequence  
Nucleic Acid Accession #: NM\_003474.2  
Coding sequence: 37..3036

20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
CACTAACGCT	CTTCCTAGTC	CCCGGGCCAA	CTCGGACAGT	TTGCTCATTT	ATTGCAACGG	60
TCAAGGCTGG	CTTGTCGCCAG	AACGGCGCGC	GCGCGACGCA	CGCACACACA	CGGGGGGAAA	120
CTTTTTTAA	AATGAAAGGC	TAGAAGAGCT	CAGCGGCGGC	GCGGGCCGTG	CGCGAGGGCT	180
CCGGAGCTGA	CTCGCCGAGG	CAGGAAATCC	CTCCGGTCGC	GACGCCCGGC	CCCCTCGGC	240
GCCCGCGTGG	GATGGTGCG	CGCTCGCCGC	CGGGCCCGAG	AGCTGCTGCA	CTGAAGGCCG	300
GCGACGATGG	CAGCGCGCCC	GCTGCCCCGTG	TCCCCCGCCC	GCGCCCTCCT	GCTCGCCCTG	360
GCCGGTGCTC	TGCTCGCGCC	CTGCGAGGCC	CGAGGGGTGA	GCTTATGGAA	CGAAGGAAGA	420
GCTGATGAAG	TTGTCACTGC	CTCTGTTCCG	AGTGGGGACC	TCTGGATCCC	AGTGAAGAGC	480
TTGACTCCA	AGAATCATCC	AGAAGTGCTG	AATATTCGAC	TACAACGGGA	AAGCAAAGAA	540
CTGATCATAA	ATCTGAAAG	AAATGAAGCT	CTCATTGCCA	GCAGTTTAC	GGAAACCCAC	600
TATCTGCAAG	ACCGTACTGA	TGTCTCCCTC	GCTCGAAAT	ACACGGTAAT	TCTGGGTAC	660
TGTTACTACC	ATGGACATGT	ACGGGGATAT	TCTGATTGAG	CAGTCAGTCT	CAGCACGTGT	720
TCTGGTCTCA	GGGGACTTAT	TGTGTTTGAA	AATGAAAGCT	ATGTCTTAGA	ACCAATGAAA	780
AGTGCAACCA	ACAGATACAA	ACTCTTCCCA	GCGAAGAAGC	TGAAAAGCGT	CCGGGGATCA	840
TGTGGATCAC	ATCACAACAC	ACCAAACCTC	GCTGCAAAGA	ATGTGTTTCC	ACCACCTCT	900
CAGACATGGG	CAAGAAGGCA	TAAAAGAGAG	ACCTCAAGG	CAACTAAGTA	TGTGGAGCTG	960
GTGATCGTGG	CAGACAACCG	AGGCAAGGAA	AGGCAAGGAA	AAGATCTGGA	AAAAGTTAAG	1020
CAGCGATTAA	TAGAGATTGC	TAATCACGTT	GACAAGTTT	ACAGACCACT	GAACATTCGG	1080
ATCGTGTGG	TAGGCGTGGA	AGTGTGGAAT	GACATGGACA	AATGCTCTGT	AAGTCAGGAC	1140
CCATTCACCA	GCCTCCATGA	ATTTCTGGAC	TGGAGGAAGA	TGAAGCTTCT	ACCTCGCAA	1200
TCCCATGACA	ATGCGCAGCT	TGTCACTGGG	GTTTATTTCC	AAGGGACCAC	CATCGGCATG	1260
GCCCCAATCA	TGAGCATGTG	CACGGCAGAC	CAGTCTGGGG	GAATTGTCT	GGACCATTC	1320
GACAATCCCC	TTGGTGACG	CGTGACCTG	GACATGAGC	TGGGCCACAA	TTTCGGGATG	1380
AATCATGACA	CACCTGGACG	GGGCTGTAGC	TGTCAAATGG	CGGTTGAGAA	AGGAGGCTGC	1440
ATCATGAACG	CTTCCACCGG	GTACCCATTT	CCCATGGTGT	TCAGCAGTTG	CAGCAGGAAG	1500
GACTTGGAGA	CCAGCCTGGA	GAAAGGAATG	GGGGTGTGCC	TGTTTAACCT	GCCGGAAGTC	1560
AGGGAGTCTT	TCCGGGGCCA	GAAAGTGTGG	AACAGATTG	TGGAAGAAGG	AGAGGAGTGT	1620
GACTGTGGGG	AGCCAGAGGA	ATGTATGAAT	CGCTGCTGCA	ATGCCACCAC	CTGTACCCTG	1680
AAGCCGGAGC	CTGTGTGCGC	ACATGGGCTG	TGCTGTGAAG	ACTGCCAGCT	GAAGCCTGCA	1740
GGAACAGCGT	GCAAGGACTC	CAGCAACTCC	TGTGACCTCC	CAGAGTTCTG	CACAGGGGCC	1800
AGCCCTCACT	GCCAGCCAA	CGTGACCTG	CACGATGGGC	ACTCATGTCA	GGATGTGGAC	1860
GGCTACTGCT	ACAATGGCAT	CTGCCAGACT	CACGAGCAGC	AGTGTGTAC	ACTCTGGGGA	1920
CCAGGTGCTA	AACCTGCCCC	TGGGATCTGC	TTTGAGAGAG	TCAATTCTGC	AGGTGATCCT	1980
TATGGCAACT	GTGGCAAAGT	CTCGAAGAGT	TCCTTTGCCA	AATGCGAGAT	GAGAGATGCT	2040
AAATGTGGAA	AAATCCAGTG	TCAAGGAGGT	GCCAGCCGGC	CAGTCATTGG	TACCAATGCC	2100
GTTTCCATAG	AAACAAACAT	CCCCCTTCAG	CAAGGAGGCC	GGATTCTGTG	CCGGGGGACC	2160
CACGTGTACT	TGGGCGATGA	CATGCCGGAC	CCAGGGCTTG	TGCTTGACAG	CACAAAGTGT	2220
GCAGATGGAA	AAATCTGCCT	GAATCGTCAA	TGTCAAATA	TTAGTGTCTT	TGGGGTTTAC	2280
GAGTGTGCAA	TGCAAGTCCA	CGGCAGAGGG	GTGTGCAACA	ACAGGAAGAA	CTGCCACTGC	2340
GAGGCCCACT	GGGCACCTCC	CTTCTGTGAC	AAGTTTGGCT	TTGGAGGAAG	CACAGACAGC	2400
GGCCCCATCC	GGCAAGCAGA	TAACCAAGGT	TTAACCATAG	GAATTCTGGT	GACCATCCTG	2460
TGTCTTCTTG	CTGCCGATTT	TGTGGTTTAT	CTCAAAAGGA	AGACCTTGAT	ACGACTGCTG	2520
TTTACAAATA	AGAAAGACCAC	CATTGAAAAA	CTAAGGTGTG	TGCGCCCTTC	CCGGCCACCC	2580
CGTGGCTTCC	AACCTGTCTA	GGCTCACCTC	GGCCACCTTG	GAAAAGGCCT	GATGAGGAAG	2640
CCGCCAGATT	CCTACCCACC	GAAGGACAAT	CCCAGGAGAT	TGCTGCAGTG	TCAGAATGTT	2700
GACATCAGCA	GACCCCTCAA	CGGCCTGAAT	GTCCCTCAGC	CCCAGTCAAC	TCAGCGAGTG	2760
CTTCTCTCCC	TCCACCGGGC	CCCACGTGCA	CCTAGCGTCC	CTGCCAGACC	CCTGCCAGCC	2820
AAGCCTGCAC	TTAGGCAGGC	CCAGGGGACC	TGTAAGCCAA	ACCCCTCTCA	GAAGCCTCTG	2880
CCTGCAGATC	CTCTGGCCAG	AACAACCTCG	CTCACTCATG	CCTTGGCCAG	GACCCAGGA	2940
CAATGGGAGA	CTGGGCTCCG	CCTGGCACCC	CTCAGACCTG	CTCCACAATA	TCCACACCAA	3000
GTGCCCAGAT	CCACCCACAC	CGCCTATATT	AAGTGAGAAG	CCGACACCTT	TTTTCAACAG	3060
TGAAGACAGA	AGTTTGCACT	ATCTTTTACG	TCCAGTTGGA	GTTTTTTGTA	CCAACTTTTA	3120
GGATTTTTTT	TAATGTTTAA	AACATCATTA	CTATAAGAAC	TTTGAGCTAC	TGCCGTCAGT	3180
GCTGTGCTGT	GCTATGGTGC	TCTGTCTACT	TGCACAGGTA	CTTGTAATTT	ATTAATTTAT	3240
GCAGAATGTT	GATTACAGTG	CAGTGCGCTG	TAGTAGGCAT	TTTTACCATC	ACTGAGTTTT	3300
CCATGGCAGG	AAGGCTTGTT	GTGCTTTTAG	TATTTTAGTG	AACTTGAAAT	ATCCTGCTTG	3360
ATGGGATTCT	GGACAGGATG	TGTTTGCTTT	CTGATCAAGG	CCTTATTGGA	AAGCAGTCCC	3420
CCAACATCCC	CCAGCTGTGC	TTATGGTACC	AGATGCAGCT	CAAGAGATCC	CAAGTAGAAT	3480
CTCAGTTGAT	TTTCTGGATT	CCCCATCTCA	GGCCAGAGCC	AAGGGGCTTC	AGGTCCAGGC	3540
TGTGTTTGGC	TTTCAGGGAG	GCCCTGTGCC	CCTTGACAAC	TGGCAGGCAG	GCTCCAGGG	3600
ACACCTGGGA	GAAATCTGGC	TTCTGGCCAG	GAAAGCTTGG	TGAGAACCCTG	GGTTGCAGAC	3660
AGGAATCTTA	AGGTGTAGCC	ACACCAGGAT	AGAGACTGGA	ACACTAGACA	AGCCAGAAGT	3720
TGACCCTGAG	CTGACCAGCC	GTGAGCATGT	TTGGAAGGGG	TCTGTAGTGT	CACTCAAGGC	3780
GGTGCTTGAT	AGAAATGCCA	AGCACTTCTT	TTTCTCGCTG	TCCTTTCTAG	AGCACTGCCA	3840



	CCAGTAGGTT	ATTTAGCTTG	GGAAAGGTGG	TGTTTCTGTA	AGAAACCTAC	TGCCCAGGCA	3900
	CTGCAAACCG	CCACCTCCCT	ATACTGCTTG	GAGCTGAGCA	AATCACCACA	AACTGTAATA	3960
	CAATGATCCT	GTATTCAGAC	AGATGAGGAC	TTTCCATGGG	ACCACAACCTA	TTTTCAGATG	4020
5	TGAACCATTA	ACCAGATCTA	GTCAATCAAG	TCTGTTTACT	GCAAGGTTCA	ACTTATTAAC	4080
	AATTAGGCAG	ACTCTTTATG	CTTGCAAAAA	CTACAACCAA	TGGAATGTGA	TGTTTCATGGG	4140
	TATAGTTCAT	GTCTGCTATC	ATTATTCGTA	GATATTGGAC	AAAGAACCTT	CTCTATGGGG	4200
	CATCCTCTTT	TTCCAACCTG	GCTGCAGGAA	TCTTTAAAAG	ATGCTTTTAA	CAGAGTCTGA	4260
	ACCTATTTCT	TAAACACTTG	CAACCTACCT	GTGAGCATC	ACAGAATGTG	ATAAGGAAAT	4320
10	CAACTTGCTT	ATCAACTTCC	TAAATATTAT	GAGATGTGGC	TTGGGCAGCA	TCCCCTTGAA	4380
	CTCTTCACTC	TTCAAATGCC	TGACTAGGGA	GCCATGTTTC	ACAAGGTCTT	TAAAGTGAAT	4440
	AATGGCATGA	GAAATACAAA	AATACTCAGA	TAAGGTAAAA	TGCCATGATG	CCTCTGTCTT	4500
	CTGGACTGGT	TTTCACATTA	GAAGACAATT	GACAACAGTT	ACATAATTCA	CTCTGAGTGT	4560
	TTTATGAGAA	AGCCTTCTTT	TGGGGTCAAC	AGTTTTCTTA	TGCTTTGAAA	CAGAAAAATA	4620
15	TGTACCAAGA	ATCTTGGTTT	GCCTTCCAGA	AAACAAAAC	GCATTTCACT	TTCCCGGTGT	4680
	TCCCCACTGT	ATCTAGGCAA	CATAGTATTC	ATGACTATGG	ATAAACTAAA	CACGTGACAC	4740
	AAACACACAC	AAAAGGGAAC	CCAGCTCTAA	TACATTCCAA	CTCGTATAGC	ATGCATCTGT	4800
	TTATTCTATA	GTTATTAAAGT	TCTTTAAAAT	GTAAAGCCAT	GCTGGAAAAT	AATACTGCTG	4860
	AGATACATAC	AGAATTACTG	TAAGTATGTA	CACTTGGTAA	TTGTACTAAA	GCCAAACATA	4920
20	TATATACTAT	TAAAAAGGTT	TACAGAATTT	TATGGTGCAT	TACGTGGGCA	TTGTCTTTTT	4980
	AGATGCCCAA	ATCCTTAGAT	CTGGCATGTT	AGCCCTTCCT	CCAATTATAA	GAGGATATGA	5040
	ACCAAAAAAA	AAAAAAAAAA	AA				

Seq ID NO: 427 Protein sequence  
Protein Accession #: NP\_003465

25	1	11	21	31	41	51	
	MAARPLPVSP	ARALLLALAG	ALLAPCEARG	VSLWNEGRAD	EVVSASVRSR	DLWIPVKSFD	60
30	SKNHPEVLNI	RLQRESKELI	INLERNEGLI	ASSFTETHYL	QDGTDVSLAR	NYTVILGHY	120
	YHGHVRGYS	SAVSLSTCSG	LRGLIVFENE	SYVLEPMKSA	TNRYKLFPK	KLKSVRGSCG	180
	SHHNTPNLAA	KNVFPSPSQT	WARRHKRETL	KATKYVELVI	VADNREFQRQ	GKDLEKVKQR	240
	LIEIANHVDK	FYRPLNIRIV	LVGVEVWVND	DKCSVSQDPF	TSLHEFLDWR	KMKLLPRKSH	300
	DNAQLVSGVY	FQGTITGMAP	IMSMCTADQS	GGIVMDHSDN	PLGAAVTLAH	ELGHNFGMNH	360
35	DTLDRGCSCQ	MAVEKGGCIM	NASTGYFPFM	VFSSCSRKDL	ETSLEKGMGV	CLFNLPEVRE	420
	SFGGQKCGNR	FVEEGEECD	GEPEECMNR	CNATTCTLKP	DAVCAHGLCC	EDCQLKPAGT	480
	ACRDSNSCD	LPEFCTGASP	HCPANVYLHD	GHSCQDVG	CYNGICQTHE	QQCVTLWPG	540
	AKPAPGICFE	RVNSAGDPY	NCCKVSKSSF	AKCEMRDAK	GKIQCGGAS	RPVIGTNAV	600
	IETNIPQLQG	GRILCRGTHV	YLGDDMPDP	LVLAGTKCAD	GKICLNRCQ	NISVFGVHE	660
40	AMQCHGRGVC	NNRKNCHCEA	HWAPFFCDKF	GFGGSTDSGP	IRQADNQGLT	IGILVTILCL	720
	LAAGFVVYLK	RKTLIRLLFT	NKKTIEKLR	CVRPSRPPRG	FQPCQAHLG	LKGGLMRKPP	780
	DSYPPKDNPR	RLQCCQNVDI	SRPLNGLNRP	QRPSTQVLP	PLHRAPRAPS	VPARPLPAKP	840
	ALRQAQGTCK	PNPPQKPLPA	DPLARTTRLT	HALARTPGQW	ETGLRLAPLR	PAPQYPHQVP	900
	RSTHTAYIK						

Seq ID NO: 428 DNA sequence  
Nucleic Acid Accession #: NM\_003714  
Coding sequence: 135..1043

50	1	11	21	31	41	51	
	GAGGAGGAGG	GAAAAGGCGA	GCAAAAAGGA	AGAGTGGGAG	GAGGAGGGGA	AGCGGCGAAG	60
	GAGGAAGAGG	AGGAGGAGGA	AGAGGGGAGC	ACAAAGGATC	CAGGTCTCCC	GACGGGAGGT	120
	TAATACCAAG	AACCATGTGT	GCCGAGCGGG	TGGCCAGATT	CATGACCCTG	GCTTTGGTGT	180
55	TGGCCACCTT	TGACCCGGCG	CGGGGACCG	ACGCCACCAA	CCCACCCGAG	GGTCCCCAAG	240
	ACAGGAGCTC	CCAGCAGAAA	GGCCGCCTGT	CCCTGCAGAA	TACAGCGGAG	ATCCAGCACT	300
	GTTTGGTCAA	CGCTGGCGAT	GTGGGTGTG	GCGTGTGTTG	ATGTTTCGAG	AACAACTCTT	360
	GTGAGATTCT	GGGCTTACAT	GGGATTTGCA	TGACTTTTCT	GCACAACGCT	GGAAAATTTG	420
	ATGCCCAGGG	CAAGTCATTC	ATCAAGACG	CCTTGAAATG	TAAGGCCAC	GCTCTGCGGC	480
60	ACAGGTTCCG	CTGCATAAGC	CGGAAGTGCC	CGGCCATCAG	GGAAATGGTG	TCCAGTTGTC	540
	AGCGGGAATG	CTACCTCAAG	CACGACCTGT	CGCGGGCTGC	CCAGGAGAAC	ACCCGGGTGA	600
	TAGTGAGAGT	GATCCATTTC	AAGGACTTGC	TGCTGCACGA	ACCCTACGTG	GACCTCGTGA	660
	ACTTGCTGCT	GACCTGTGGG	GAGGAGGTGA	AGGAGGCCAT	CACCCACAGC	GTGCAGGTTC	720
	AGTGTGAGCA	GAAGTGGGGA	AGCCTGTGCT	CCATCTTGAG	CTTCTGCACC	TCCGCCATCC	780
	AGAAGCCTCC	CACGGCGCCC	CCCGAGCGCC	AGCCCCAGGT	GGACAGAAC	AAGCTCTCCA	840
65	GGGCCCACCA	CGGGGAAGCA	GGACATCACC	TCCAGAGCC	CAGCAGTAGG	GAGACTGGCC	900
	GAGGTGCCAA	GGGTGAGCGA	GGTAGCAAGA	GCCACCCAAA	CGCCCATGCC	CGAGGCAGAG	960
	TCGGGGGCTT	TGGGGCTCAG	GGACCTTCCG	GAAGCAGCGA	GTGGGAAGAC	GAACAGTCTG	1020
	AGTATTCTGA	TATCCGGAGG	TGAAATGAAA	GGCCTGGCCA	CGAAATCTTT	CCTCCACGCC	1080
70	GTCCATTTTC	TTATCTATGG	ACATTCCTAA	ACATTTACCA	TTAGAGAGGG	GGGATGTGAC	1140
	ACGCAGGATT	CTGTGGGGAC	TGTGGACTTC	ATCGAGGTGT	GTGTTTCGCG	AACGGACAGG	1200
	TGAGATGGAG	ACCCCTGGGG	CCGTGGGGTC	TCAGGGGTGC	CTGGTGAATT	CTGCACTTAC	1260
	ACGTACTCAA	GGGAGCGCGC	CCGCTTATC	CTCGTACCTT	TGTCTTCTTT	CCATCTGTGG	1320
	AGTCAGTGGG	TGTCCGGCCG	TCTGTTGTGG	GGGAGGTGAA	CCAGGGAGGG	GCAGGGCAAG	1380
75	GCAGGGCCCC	CAGAGCTGGG	CCACACAGTG	GGTGTGGGGC	CTCGCCCCGA	AGCTTCTGGT	1440
	GCAGCAGCCT	CTGGTGCTGT	CTCCGCGGAA	GTCAGGGCGG	CTGGATTCCA	GGACAGGAGT	1500
	GAATGTAAAA	ATAAATATCG	CTTAGAATGC	AGGAGAAGGG	TGGAGAGGAG	GCAGGGGCCG	1560
	AGGGGGTGCT	TGGTGCCAAA	CTGAAATTC	GTTTCTTGTG	TGGGGCCTTG	CGGTTGAGAG	1620
	CTCTTGGCGA	GGGTGGAGGG	AGGAGTGTG	TTTCTATGTG	TAATTTCTGA	GCCATTGTAC	1680
80	TGTCTGGGCT	GGGGGGGACA	CTGTCCAAGG	GAGTGGCCCC	TATGAGTTTA	TATTTTAACC	1740
	ACTGCTTCAA	ATCTCGATTT	CACCTTTTTT	ATTTATCCAG	TTATATCTAC	ATATCTGTCA	1800
	TCTAAATAAA	TGGCTTTCAA	ACAAAGCAAC	TGGGTGATTA	AAACCAGCTC	AAAGGGGGTT	1860
	TAAAAAATAA	AAAACAGGCC	CATCCTTTGA	GGCTGATTTT	TCTTTTTTTT	AAGTTCTATT	1920
	TTAAAAGCTA	TCAAACAGCG	ACATAGCCAT	ACATCTGACT	GCCTGACATG	GACTCCTGCC	1980
	CACTTGGGGG	AAACCTTATA	CCCAGAGGAA	AATACACACC	TGGGGAGTAC	ATTGACAAA	2040
85	TTCCCTTAG	GATTTCTGTA	TCTCACCTTG	ACCTCAGGCC	AAGATTGGTA	AAGCTGCGTC	2100
	CTGGCGATTC	CAGGAGACCC	AGCTGGAAAC	CTGGCTTCTC	CATGTGAGGG	GATGGGAAAG	2160
	GAAAGAAGAG	AATGAAGACT	ACTTAGTAAT	TCCCATCAGG	AAATGCTGAC	CTTTTACATA	2220



AAATCAAGGA GACTGCTGAA AATCTCTAAG GGACAGGATT TTCCAGATCC TAATTGGAAA 2280  
 TTTAGCAATA AGGAGAGGAG TCCAAGGGGA CAAATAAAGG CAGAGAGAGA GAGAGAGAGA 2340  
 GGGAGAGGAA GAAAAGAGAG AGAGAAAAGA GCCTCGTGCC

5 Seq ID NO: 429 Protein sequence  
 Protein Accession #: NP\_003705

10 1 11 21 31 41 51  
 | | | | | |  
 MCAERLGQFM TLALVLATFD PARGTDATNP PEGPQDRSSQ QKGRLSLQNT AEIQHCLVNA 60  
 GDVGCQVFEC FENNSCEIRG LHGICMTFLH NAGKFDAQGK SFIKDALKCK AHALRHRFGC 120  
 ISRKCPAIRE MVSQLQRECY LKHDLCAAAQ ENTRVIVEMI HFKDLLLHEP YVDLVNLLLT 180  
 CGEEVKEAIT HSVQVQCEQN WGSLSLSILSF CTSIAIQKPPT APPERQPOVD RTKLSRAHHG 240  
 EAGHHLPEPS SRETGRGAKG ERGSKSHPNA HARGRVGGLG AQGPSGSSEW EDEQSEYSDI 300  
 RR

Seq ID NO: 430 DNA sequence  
 Nucleic Acid Accession #: NM\_005940  
 Coding sequence: 23..1489

20 1 11 21 31 41 51  
 | | | | | |  
 AAGCCCAGCA GCCCCGGGGC GGATGGCTCC GGCCGCCTGG CTCCGCAGCG CGGCCGCGCG 60  
 CGCCCTCCTG CCCCCGATGC TGCTGTGTGCT GCTCCAGCCG CCGCCGCTGC TGGCCCGGGC 120  
 25 TCTGCCGCGG GACGTCCACC ACCTCCATGC CGAGAGGAGG GGGCCACAGC CCTGGCATGC 180  
 AGCCCTGCCC AGTAGCCCGG CACCTGCCCC TGCCACGCAG GAAGCCCCC GGCCTGCCAG 240  
 CAGCCTCAGG CCTCCCCGCT GTGGCGTGCC CGACCCATCT GATGGGCTGA GTGCCCGCAA 300  
 CCGACAGAAG AGGTTCGTGC TTTCTGGCGG GCGCTGGGAG AAGACGGACC TCACCTACAG 360  
 GATCCTTCGG TTTCCATGGC AGTTGGTGCA GGAGCAGGTG CGGCAGACGA TGGCAGAGGC 420  
 30 CCTAAAGGTA TGGAGCGATG TGACGCCACT CACCTTTACT GAGGTGCACG AGGGCCGTGC 480  
 TGACATCATG ATCGACTTCG CCAGGTACTG GCATGGGGAC GACCTGCCGT TTGATGGGCC 540  
 TGGGGGCATC CTGGCCCATG CCTTCTTCCC CAAGACTCAC CGAGAAGGGG ATGTCCACTT 600  
 CGACTATGAT GAGACCTGGA CTATCGGGGA TGACCAGGGC ACAGACCTGC TGCAGGTGGC 660  
 AGCCCATGAA TTTGGCCACG TGCTGGGGCT GCAGCACACA ACAGCAGCCA AGGCCCTGAT 720  
 35 GTCCGCCTTC TACACCTTTC GCTACCCACT GAGTCTCAGC CCAGATGACT GCAGGGCGGT 780  
 TCAACACCTA TATGGCCAGC CCTGGCCAC TGTCACCTCC AGGACCCAG CCCTGGGCCC 840  
 CCAGGCTGGG ATAGACACCA ATGAGATTGC ACCGCTGGAG CCAGACGCC CGCCAGATGC 900  
 CTGTGAGGCC TCCTTTGACG CGGTCTCCAC CATCCGAGGC GAGCTCTTTT TCTTCAAAGC 960  
 GGGCTTTGTG TGGCGCCTCC GTGGGGCCCA GCTGCAGCCC GGCTACCCAG CATTGGCCTC 1020  
 40 TCGCCACTGG CAGGGACTGC CCAGCCCTGT GGACGCTGCC TTCGAGGATG CCCAGGGCCA 1080  
 CATTGTGTTT TTCCAAGGTG CTCAGTACTG GGTGTACGAC GGTGAAAAGC CAGTCTGGG 1140  
 CCCCACACCC CTCACCGAGC TGGGCCTGGT GAGGTTCCTG GTCCATGCTG CCTTGGTCTG 1200  
 GGGTCCCAG AAGAACAAGA TCTACTTCTT CCGAGGCAGG GACTACTGGC GTTTCCACCC 1260  
 CAGCACCCGG CGTGTAGACA GTCCCGTGCC CCGCAGGGCC ACTGACTGGA GAGGGGTGCC 1320  
 45 CTCTGAGATC GACGCTGCCT TCCAGGATGC TGATGGCTAT GCCTACTTCC TGC CGCGGCCG 1380  
 CCTCTACTGG AAGTTTGACC CTGTGAAGTC GAAGGCTCTG GAAGGCTTCC CCCGTCTCGT 1440  
 GGGTCTTGAC TTCTTTGGCT GTGCCGAGCC TGCCAACACT TTCCTCTGAC CATGGCTTGG 1500  
 ATGCCCTCAG GGGTGCTGAC CCCTGCCAGG CCACGAATAT CAGGCTAGAG ACCCATGGCC 1560  
 50 ATCTTTGTGG CTGTGGGCAC CAGGCATGGG ACTGAGCCCA TGTCTCTGTC AGGGGGATGG 1620  
 GGTGGGGTAC AACCACCATG ACAACTGCCG GGAGGGCCAC GCAGGTCTGT GTACCTGCC 1680  
 AGCGACTGTC TCAGACTGGG CAGGGAGGCT TTGGCATGAC TTAAGAGGAA GGGCAGTCTT 1740  
 GGGACCCGCT ATGCAGGTCC TGGCAAACCT GGCTGCCCTG TCTCATCCCT GTCCCTCAGG 1800  
 GTAGCACCAT GGCAGGACTG GGGGAACCTG AGTGTCTTGT CTGTATCCCT GTTGTGAGGT 1860  
 55 TCCTTCCAGG GGCTGGCACT GAAGCAAGGG TGCTGGGGCC CCATGGCCTT CAGCCCTGGC 1920  
 TGAGCAACTG GGCTGTAGGG CAGGGCCACT TCCTGAGGTC AGGTCTTGGT AGGTGCCTGC 1980  
 ATCTGTCTGC CTTCTGGCTG ACAATCCTGG AAATCTGTTC TCCAGAATCC AGGCCAAAAA 2040  
 GTTACAGTCA AAATGGGGAG GGGTATTCTT CATGCAGGAG ACCCCAGGCC CTGGAGGCTG 2100  
 CAACATACCT CAATCCTGTC CCAGGCCGGA TCCTCCTGAA GCCCTTTTCG CAGCACTGCT 2160  
 60 ATCCTCCAAA GCATTGTAA ATGTGTGTAC AGTGTGTATA AACCTTCTTC TTCTTTTTTT 2220  
 TTTTAAACT GAGGATTGTC ATTAAACACA GTTGTTTTCT

Seq ID NO: 431 Protein sequence  
 Protein Accession #: NP\_005931

65 1 11 21 31 41 51  
 | | | | | |  
 MAPAAWLRSA AARALLPPML LLLLQPPPLL ARALPPDVHH LHAERRGPQP WHAALPSSPA 60  
 PAPATQEAPR PASLRPPRC GVPDPSDGLS ARNRQKRFVL SGRWEKTDL TYRILRFPWQ 120  
 70 LVQEQVRQTM AEALKVWSDV TPLTFTEVHE GRADIMIDFA RYWHGDDLFP DPGGILAHA 180  
 FFPKTHREGD VHFYDETWT IGDDQGTDL QVAAHEFGHV LGLQHTTAAK ALMSAFYTFR 240  
 YPLSLSPDDC RGVQHLYGQP WPTVTSRTPA LGPQAGIDTN EIAPLEPDAP PDACEASFDA 300  
 VSTIRGELFF FKAGFVWRLR GGQLQPGYPA LASRHWQGLP SPVDAAFEDA QGHIWFFQGA 360  
 QYWVYDGEKP VLGPAPLTEL GLVRFPVHAA LVWGPEKNKI YFFRGRDYWR FHPSTRRVD 420  
 75 PVPRRATDWR GVPSEIDAAF QDADGYAYFL RGRLYWKFDV VKVKALEGFP RLVGPDFFGC 480  
 AEPANTFL

Seq ID NO: 432 DNA sequence  
 Nucleic Acid Accession #: NM\_024022  
 Coding sequence: 202..1563

80 1 11 21 31 41 51  
 | | | | | |  
 ACCGGGCACC GGACGGCTCG GGTACTTTCG TTCTTAATTA GGTCATGCCC GTGTGAGCCA 60  
 GGAAAGGGCT GTGTTTATGG GAAGCCAGTA ACACGTGTGGC CTACTATCTC TTCCGTGGTG 120  
 85 CCATCTACAT TTTTGGGACT CGGGAATTAT GAGGTAGAGG TGGAGGCGGA GCCGGATGTC 180  
 AGAGGTCCTG AAATAGTCAC CATGGGGGAA AATGATCCGC CTGCTGTTGA AGCCCCCTTC 240  
 TCATTCCGAT CGCTTTTGG CCTTGATGAT TTGAAAATAA GTCTGTGTC ACCAGATGCA 300



5 GATGCTGTTG CTGCACAGAT CCTGTCACTG CTGCCATTGA AGTTTTTTCC AATCATCGTC 360  
 ATTGGGATCA TTGCATTGAT ATTAGCACTG GCCATTGGTC TGGGCATCCA CTTGCACTGC 420  
 TCAGGGAAGT ACAGATGTCG CTCATCCTTT AAGTGTATCG AGCTGATAGC TCGATGTGAC 480  
 GGAGTCTCGG ATTGCAAAGA CGGGGAGGAC GAGTACCGCT GTGTCCGGGT GGGTGGTCAG 540  
 AATGCCGTGC TCCAGGTGTT CACAGCTGCT TCGTGGAAGA CCATGTGCTC CGATGACTGG 600  
 AAGGGTCACT ACGCAAATGT TGCCTGTGCC CAACTGGGTT TCCCAAGCTA TGTGAGTTCA 660  
 GATAACCTCA GAGTGAGCTC GCTGGAGGGG CAGTTCCGGG AGGAGTTTGT GTCCATCGAT 720  
 CACCTCTTGC CAGATGACAA GGTGACTGCA TTACACCACT CAGTATATGT GAGGGAGGGA 780  
 TGTECCTCTG GCCACGTGGT TACCTTGCA GTCACAGCCT GTGGTCATAG AAGGGGCTAC 840  
 10 AGCTCACGCA TCGTGGGTGG AAACATGTCC TTGCTCTCGC AGTGGCCCTG GCAGGCCAGC 900  
 CTTCACTTCC AGGGCTACCA CCTGTGCGGG GGCTCTGTCA TCACGCCCTT GTGGATCATC 960  
 ACTGCTGCAC ACTGTGTTTA TGACTTGTA CTCCCCAAGT CATGGACCAT CCAGGTGGGT 1020  
 CTAGTTTCCC TGTTGGACAA TCCAGCCCCA TCCCACTTGG TGGAGAAGAT TGTCTACCAC 1080  
 15 AGCAAAGTACA AGCCAAAGAG GCTGGGCAAT GACATCGCCC TTATGAAGCT GGCCGGGCCA 1140  
 CTCACGTTC AATGAAATGAT CCAGCCTGTG TGCCTGCCCA ACTCTGAAGA GAACTTCCCC 1200  
 GATGGAAGAG TGTGCTGGAC GTCAGGATGG GGGGCCACAG AGGATGGAGG TGACGCCTCC 1260  
 CCTGTCTGTA ACCACGCGGC CGTCCCTTTG ATTTCCAACA AGATCTGCAA CCACAGGGAC 1320  
 GTGTACGGTG GCATCATCTC CCCCTCCATG CTCTGCGCGG GCTACCTGAC GGGTGGCGTG 1380  
 20 CAGAGCTGCC AGGGGGACAG CGGGGGGCCC CTGGTGTGTC AAGAGAGGAG GCTGTGGAAG 1440  
 TTAGTGGGAG CGACCAAGCTT TGGCATCGGG TGCGCAGAGG TGAACAAGCC TGGGGTGTAC 1500  
 ACCCGTGTCA CCTCCTTCC TGGACTGGATC CACGAGCAGA TGGAGAGAGA CCTAAAAACC 1560  
 TGAAGAGGAA GGGGACAAGT AGCCACCTGA GTTCTGAGG TGATGAAGAC AGCCCGATCC 1620  
 TCCCTGGAC TCCCGTGTAG GAACCTGCAC ACGAGCAGAC ACCCTTGGAG CTCTGAGTTC 1680  
 CGGCACCAAGT AGCAGGCCCG AAAGAGGCAC CCTTCCATCT GATTCCAGCA CAACCTTCAA 1740  
 25 GCTGCTTTTT GTTTTTTGT TTTTGGAGT GGAGTCTCGC TCTGTTGCC AGGCTGGAGT 1800  
 GCAGTGGCGA AATCCTGTCT CACTGCAGCC TCCGCTTCCC TGGTCAAGC GATTCTCTTG 1860  
 CCTCAGCTTC CCCAGTAGCT GGGACCACAG GTGCCCCGCA CCACACCCAA CTAATTTTGT 1920  
 TATTTTTAGT AGAGACAGGG TTTCACCATG TTGGCCAGGC TGCTCTCAA CCCCTGACCT 1980  
 30 CAAATGATGT GCCTGCTTCA GCCTCCACA GTGCTGGGAT TACAGGCATG GGCCACCACG 2040  
 CCTAGCCTCA CGCTCCTTTC TGATCTTAC TAAGAACAAA AGAAGCAGCA ACTTGCAAGG 2100  
 GCGGCCTTTC CCACTGGTCC ATCTGGTTTT CTCTCCAGGG GTCTTGCAA ATTCTGACG 2160  
 AGATAAGCAG TTATGTGACC TCACGTGCAA AGCCACCAAC AGCCACTCAG AAAAGACGCA 2220  
 CCAGCCCAAG AGTGCAAGAC TGCAGTCTC ATCTCTAGGG ACCAGAACCA 2280  
 35 AACCACCCCT TTCTACTTCC AAGACTTATT TTCACATGTG GGGAGGTTAA TCTAGGAATG 2340  
 ACTCGTTTAA GGCCTATTTT CATGATTTCT TTGTAGCATT TGGTGCTTGA CGTATTATTG 2400  
 TCCTTTGATT CCAAATAATA TGTTCCTTC CCTCAAAAAA AAAAAA AAAAAA 2460  
 AAAAA

Seq ID NO: 433 Protein sequence  
 Protein Accession #: NP\_076927

45 1 11 21 31 41 51  
 | | | | | |  
 MGENDPPAVE APFSFRSLFG LDDLKISFVA PDADAVAQI LSLPLKFFP IIVIGIIALI 60  
 LALAIGLGIH FDCSGKYRCR SFFKCIELIA RCDGVSDCKD GEDEYRCVRV GGQNAVLQVF 120  
 TAASWKTMCS DDWKGHYANV ACAQLGFPSY VSSDNLRVSS LEGQFREEFV SIDHLLPDDK 180  
 VTALHHSVYV REGCAGSHV TLQCTACGHR RGYSSRIVGG NMSLLSQWPW QASLQFQGYH 240  
 LCGGSVITPL WIITAAHCVY DLYLPKSWTI QVGLVSLLDN PAPSHLVEKI VYHSKYKPKR 300  
 50 LGNDIALMKL AGPLTFNEMI QPVCLPNSEE NFPDGKVCWT SGWGATEDGG DASPVLNHAA 360  
 VPLISNKICN HRDVYGGIIS PSMCLCAGYLT GGVDSCQGDS GGPLVCQERR LWKLVGATSF 420  
 GIGCAEVNKP GVIYTRVTSFL DWIHEQMERD LKT

Seq ID NO: 434 DNA sequence  
 Nucleic Acid Accession #: NM\_000493.2  
 Coding sequence: 97..2139

55 1 11 21 31 41 51  
 | | | | | |  
 CACCTTCTGC ACTGCTCATC TGGGCAGAGG AAGCTTCAGA AAGCTGCCAA GGCAACATCT 60  
 CCAGGAACTC CCAGCACGCA GAATCCATCT GAGAATATGC TGCCACAAAT ACCCTTTTGT 120  
 CTGCTAGTAT CCTTGAACCTT GGTTTCATGGA GTGTTTTACG CTGAACGATA CCAAATGCCC 180  
 ACAGGCATAA AAGGCCCACT ACCCAACACC AAGACACAGT TCTTCATTCC CTACACCATA 240  
 AAGAGTAAAG GTATAGCAGT AAGAGGAGAG CAAGGTACTC CTGGTCCACC AGGCCCTGCT 300  
 65 GGACCTCGAG GGCACTCAGG TCCTTCTGGA CCACCAAGAA AACCAGGCTA CGGAAGTCCT 360  
 GGACTCCAAG GAGAGCCAGG GTTGCCAGGA CCACCGGGAC CATCAGCTGT AGGGAACCA 420  
 GGTGTGCCAG GACTCCCAGG AAAACCAGGA GAGAGAGGAC CATATGGACC AAAAGGAGAT 480  
 GTTGGACCAG CTGGCCTACC AGGACCCCGG GGCCACCAAG GACCACCTGG AATCCCTGGA 540  
 CCGGCTGGAA TTTCTGTGCC AGGAAAACCT GGACAACAGG GACCCACAGG AGCCCCAGGA 600  
 CCCAGGGGCT TTCCTGGAGA AAAGGGTGCA CCAGGAGTCC CTGGTATGAA TGGACAGAAA 660  
 70 GGGGAAATGG GATATGGTGC TCCTGGTCTG CCAGGTGAGA GGGGTCTTCC AGGCCCTCAG 720  
 GGTCCACAG GACCATCTGG CCTCCTGGA GTGGGAAAAA GAGGTGAAAA TGGGGTTCCA 780  
 GGACAGCCAG GCATCAAAGG TGATAGAGGT TTTCCGGGAG AAATGGGACC AATTGGCCCA 840  
 CCAGGTCCCC AAGGCCCTCC TGGGGAACGA GGGCCAGAAG GCATTGGAAG GCCAGGAGCT 900  
 75 GCTGGAGCCC CAGGCCAGCC AGGGATTCCA GGAACAAAAG GTCTCCCTGG GGCTCCAGGA 960  
 ATAGCTGGGC CCCAGGGGCC TCCTGGGCTT GCGAAACCAG GCTTGCCAGG CCTGAAGGGA 1020  
 GAAAGAGGAC CTGCTGGCCT TCCTGGGCTT CCAGGTGCCA AAGGGGAACA AGGGCCAGCA 1080  
 GGTCTTCTTG GGAAGCCAGG TCTGACTGGA CCCCCTGGGA ATATGGGACC CCAAGGACCA 1140  
 AAAGGCATCC CGGGTAGCCA TGGTCTCCCA GGCCCTAAAG GTGAGACAGG GCCAGCTGGG 1200  
 80 CCTGCAGGAT ACCCTGGGGC TAAGGGTGAA AGGGGTTCCT CTGGGTGAGA TGGAAAACCA 1260  
 GGGTACCCAG GAAAACCAGG TCTCGATGGT CCTAAGGGTA ACCAGGGTT ACCAGGTCCA 1320  
 AAAGGTGATC CTGGAGTTGG AGGACCTCCT GGTCTCCAG GCCCTGTGGG CCCAGCAGGA 1380  
 GCAAAGGGAA TGCCCGGACA CAATGGAGAG GCTGCGCCAA GAGGTGCCCC TGGAAATACCA 1440  
 GGTACTAGAG GCCCTATTGG GCCACAGGC ATTCCAGGAT TCCCTGGGTC TAAAGGGGAT 1500  
 85 CCAGGAAGTC CCGGTCTTCC TGGCCAGCT GGCATAGCAA CTAAGGGCCT CAATGGACCC 1560  
 ACCGGGCCAC CAGGGCCTCC AGGTCCAAGA GGCCACTCTG GAGAGCCTGG TCTTCCAGGG 1620  
 CCCCCTGGGC CTCCAGGCCC ACCAGGTCAA GCAGTCATGC CTGAGGGTTT TATAAAGGCA 1680  
 GGCCAAAGGC CCAGTCTTTC TGGGACCCCT CTTGTTAGTG CCAACCAGGG GGTAAACAGGA 1740



5  
10  
15  
20  
25

```
ATGCCTGTGT CTGCTTTTAC TGTATTCTC TCCAAAGCTT ACCCAGCAAT AGGAACTCCC 1800
ATACCATTTG ATAAAAATTT GTATAACAGG CAACAGCATT ATGACCCAAG GACTGGAATC 1860
TTTACTTGTC AGATACCAGG AATATACTAT TTTTCATACC ACGTGCATGT GAAAGGGACT 1920
CATGTTTGGG TAGGCCTGTA TAAGAATGGC ACCCCTGTAA TGTACACCTA TGATGAATAC 1980
ACCAAAGGCT ACCTGGATCA GGCTTCAGGG AGTGCCATCA TCGATCTCAC AGAAAATGAC 2040
CAGGTGTGGC TCCAGCTTCC CAATGCCGAG TCAAATGGCC TATACTCCTC TGAGTATGTC 2100
CACTCCTCTT TCTCAGGATT CCTAGTGGCT CCAATGTGAG TACACCCAC AGAGCTAATC 2160
TAAATCTTGT GCTAGAAAAA GCATTCTCTA ACTCTACCCC ACCCTACAAA ATGCATATGG 2220
AGGTAGGCTG AAAAGAATGT AATTTTTATT TTCTGAAATA CAGATTGAG CTATCAGACC 2280
AACAAACCTT CCCCCTGAAA AGTGAGCAGC AACGTAAAAA CGTATGTGAA GCCTCTCTTG 2340
AATTTCTAGT TAGCAATCTT AAGGCTCTTT AAGGTTTTCT CCAATATTAA AAAATATCAC 2400
CAAAGAAGTC CTGCTATGTT AAAAAACAAAC AACAAAAAAC AAAGCAACAA AAAAAAAAT 2460
TAAAAAATAA AACAGAAATA GAGCTCTAAG TTATGTGAAA TTTGATTGA GAAACTCGGC 2520
ATTTCTTTT TAAAAAAGCC TGTTCCTAAC TATGAATATG AGAACTCTA GGAAACATCC 2580
AGGAGGTATC ATATACTTT GTAGAACTTA AATACTTGAA TATTCAAAT TAAAAGACAC 2640
TGTATCCCT AAAATATTT TGATGGTGCA CTACTCTGAG GCCTGTATGG CCCCTTTCAT 2700
CAATATCTAT TCAAATATAC AGGTGCATAT ATACTTGTTA AAGCTCTTAT ATAAAAAGC 2760
CCCAAAATAT TGAAGTTCAT CTGAAATGCA AGGTGCTTTC ATCAATGAAC CTTTTCAAAA 2820
CTTTCTATG ATTGCAGAGA AGCTTTTTAT ATACCCAGCA TAAGTTGGA ACAGGTATCT 2880
GACCTATTCT TATTAGTTA ACACAAGTGT GATTAATTG ATTTCTTAA TTCCTTATTG 2940
AATCTTATGT GATATGATT TCTGGATTTA CAGAACATTA GCACATGTAC CTTGTGCCTC 3000
CCATTCAAGT GAAGTTATAA TTTACACTGA GGGTTTCAA ATTCGACTAG AAGTGGAGAT 3060
ATATTATTTA TTTATGCACT GTACTGTATT TTTATATTGC TGTTTAAAC TTTTAAGCTG 3120
TGCCTCACTT ATTAAGCAC AAAATGTTT ACCTACTCCT TATTTACGAC ACAATAAAAT 3180
AACATCAATA GATTTTtagg CTGAATTAAT TTGAAAGCAG CAATTTGCTG TTCTCAACCA 3240
TTCTTTCAAG GCTTTTCATT CGACACAATA AAATAACATC AATAG
```

Seq ID NO: 435 Protein sequence  
Protein Accession #: NP\_000484.2

30  
35  
40  
45

```
1 11 21 31 41 51
| | | | |
MLPQIPFLLL VSLNLVHGVF YAERYQMPGT IKGELPNTKT QFFIPYTIKS KGI AVRGEQG 60
TPGPPGPAGP RGHPPGSPGP GKPGYGSPL QGEPGLPGPP GPSAVGKPGV PGLPGKPGER 120
GPYGPBGDVG PAGLPGRGP PGPPGIPGPA GISVPGKPGQ QGPTGAPGR GFPGKGPAG 180
VPMNGQKGE MGYGAPGRPG ERGLPGPQGP TGPSGPPGVG KRGENGVPGQ PGIKGRGFP 240
GEMGPIGPPG PQGPPGERGP EGIGKPGAAG APGQPGIPGT KGLPGAPGIA GPPGPPGFGK 300
PGLPGLKGER GPAGLPGGPG AKGEQGPAGL PGKPLTGPP GNMGPQGPKG IPGSHGLPGP 360
KGETGPAGPA GYPGAKGERG SPGSDGKPGY PGKPLDGPB GNPGLPGPKG DPGVGGPPGL 420
PGPVGPAGAK GMPGHNGEAG PRGAPGIPGT RGPIGPPGIP GFPGSKGDPG SPGPPGPAGI 480
ATKGLNGPTG PPGPPGPRGH SGEPLPGPP GPPGPPGQAV MPEGFIKAGQ RPSLSGTPLV 540
SANQGVGTMP VSAFTVILSK AYPATGPIPI FDKILYNRQO HYDPRTGIFT CQIPGIYYFS 600
YHVHVKGTHV WVGLYKNGTP VMYTYDEYTK GYLDQASGSA IIDLTENDQV WLQLPNAESN 660
GLYSSEYVHS SFSGFLVAPM
```

Seq ID NO: 436 DNA sequence  
Nucleic Acid Accession #: XM\_062811  
Coding sequence: 1..888

50  
55  
60  
65

```
1 11 21 31 41 51
| | | | |
ATGTGGGGCG CTCGCCGCTC GTCCGTCTCC TCATCCTGGA ACGCCGCTTC GTCCTGCAG 60
CTGCTGCTGG CTGCGCTGCT GCGGCGGGG GCGAGGGCCA GCGGCGAGTA CTGCCACGGC 120
TGGCTGGACG CGCAGGGCGT CTGGCGCATC GGCTTCCAGT GTCCCGAGCG CTTCGACGGC 180
GGCGACGCCA CCATCTGCTG CCGCAGCTGC GCGTTGCGCT ACTGCTGCTC CAGCGCCGAG 240
GCGCGCCTGG ACCAGGGCGG CTGCGACAAT GACCGCCAGC AGGGCGCTGG CGAGCCTGGC 300
CGGGCGGACA AAGACGGCCC CGACGGCTCG CAGTGCCCA TCTACGTGCC GTTCCTCATT 360
GTTGGCTCCG TGTGTGTCG CTTTATCATC TTGGGTCCC TGGTGGCAGC CTGTTGCTGC 420
AGATGTCTCC GGCCTAAGCA GGATCCCCAG CAGAGCCGAG CCCCAGGGG TAACCGCTTG 480
ATGGAGACCA TCCCCTATGAT CCCCAGTGCC AGCACCTCCC GGGGGTCGTC CTCACGCCAG 540
TCCAGCACAG CTGCCAGTTC CAGCTCCAGC GCCAACTCAG GGGCCCGGGC GCCCCAACA 600
AGGTCACAGA CCAACTGTTG CTTGCCGGA GGGACCATGA ACAACGTGTA TGTCAACATG 660
CCCACGAATT TCTCTGTGCT GAACTGTCAG CAGGCCACCC AGATTGTGCC ACATCAAGGG 720
CAGTATCTGC ATCCCCATA CGTGGGTAC ACGGTGCAGC ACGACTCTGT GCCCATGACA 780
GCTGTGCCAC CTTTCATGGA CGGCCTGCAG CCTGGCTACA GGCAGATTCA GTCCCCCTTC 840
CCTCACACCA ACAGTGAACA GAAGATGTAC CCAGCGGTGA CTGTATAA
```

Seq ID NO: 437 Protein sequence  
Protein Accession #: XP\_062811

70  
75  
80

```
1 11 21 31 41 51
| | | | |
MWGARRSSVS SSWNAASLLQ LLLAALLAAG ARASGEYCHG WLDAQGVWRI GFQCPERFDG 60
GDATICCGSC ALRYCCSSAE ARLDQGGCDN DRQQGAGEPG RADKDGPDGS AVPIYVPFLI 120
VGSVFVAFII LGSIVAACCC RCLRPKQDPQ QSRAPGGNRL METIPMIPSA STSRGSSSRQ 180
SSTAASSSSS ANSGARAPPT RSQTNCLLPE GTMNNVYVNM PTNFSVLNCQ QATQIVPHQG 240
QYLHPPYVGY TVQHDSVPMT AVPPFMDGLQ PGYRQIQSPF PHTNSEQKMY PAVTV
```

Seq ID NO: 438 DNA sequence  
Nucleic Acid Accession #: NM\_004004.1  
Coding sequence: 1..681

85

```
1 11 21 31 41 51
| | | | |
ATGGATTGGG GCACGCTGCA GACGATCCTG GGGGGTGTGA ACAAACACTC CACCAGCATT 60
GGAAAGATCT GGCTACCGT CCTCTTCATT TTTTCGATTA TGATCCTCGT TGTGGCTGCA 120
AAGGAGGTGT GGGGAGATGA GCAGGCCGAC TTTGTCTGCA ACACCCTGCA GCCAGGCTGC 180
```



AAGAACGTGT GCTACGATCA CTACTTCCCC ATCTCCCACA TCCGGCTATG GGCCCTGCAG 240  
CTGATCTTTCG TGTCCAGCCC AGCGCTCCTA GTGGCCATGC ACGTGGCCTA CCGGAGACAT 300  
GAGAAGAAGA GGAAGTTCAT CAAGGGGGAG ATAAAGAGTG AATTTAAGGA CATCGAGGAG 360  
ATCAAAACCC AGAAGGTCCG CATCGAAGGC TCCCTGTGGT GGACCTACAC AAGCAGCATC 420  
TTCTTCCGGG TCATCTTCGA AGCCGCCTTC ATGTACGTCT TCTATGTCAT GTACGACGGC 480  
TTCTCCATGC AGCGGTGGT GAAGTGCAAC GCCTGGCCTT GTCCCAACAC TGTGGACTGC 540  
TTTGTGTCCC GGCCACGGA GAAGACTGTC TTCACAGTGT TCATGATTGC AGTGTCTGGA 600  
ATTTGCATCC TGCTGAATGT CACTGAATTG TGTATTTCG TAATTAGATA TTGTTCTGGG 660  
AAGTCAAAAA AGCCAGTTTA A

Seq ID NO: 439 Protein sequence  
Protein Accession #: NP\_003995.1

1 11 21 31 41 51  
| | | | |  
MDWGTLLQTL GGVNKHSTSI GKIWLTVLFI FRIMILVVAA KEVWGDEQAD FVCNTLQPGC 60  
KNVCYDHYFP ISHIRLWALQ LIFVSSPALL VAMHVAYRRH EKRRKFIKGE IKSEFKDIEE 120  
IKTQKVRIEG SLWWTYTSSI PFRVIFEAAF MYVFYVMYDG FSMQRLVKCN AWPCPNTVDC 180  
FVSRPTEKTV FTVFMIAVSG ICILLNVTEL CYLLIRYCSG KSKKPV

Seq ID NO: 440 DNA sequence  
Nucleic Acid Accession #: XM\_061091.1  
Coding sequence: 1..2481

1 11 21 31 41 51  
| | | | |  
ATGCCAAATA CTTAGGAAC AACCAGGATT GAAATTTGGC TTCTCCAAGA GCCGCCCGGG 60  
CACCGAGCGC TGGTCGCCGC TCTCCTTCCG GTGAGTCCCA GCCCCGAGTT GGCTCTGGCG 120  
CCCGGGTACC CGCCAGTGCC GGCTGCCGAT GACCGATTCA CGCTCCCGAT GATTGGAGGT 180  
CAGATGCATG GTGAGAAGGT AGATCTCTGG AGCCTTGGTG TTCTTTGCTA TGAATTTTTTA 240  
GTTGGGAAGC CTCCTTTTGA GGCAAACGAA GTCCATGTAA GCAAAGAAAC CATCGGGAAG 300  
ATTTAGCTG CCAGCAAAAT GATGTGGTGC TCGGCTGCAG TGGACATCAT GTTCTGTGTA 360  
GATGGGTCTA ACAGCGTCGG GAAAGGGAGC TTTGAAAGGT CCAAGCACTT TGCCATCACA 420  
GTCTGTGACG GTCTGGACAT CAGCCCCGAG AGGGTCAGAG TGGGAGCAAT CCAGTTCAGT 480  
TCCACTCCTC ATCTGGAATT CCCCTTGGAT TCATTTTCAA CCCAACAGGA AGTGAAGGCA 540  
AGAATCAAGA GGATGGTTTT CAAAGGAGGG CGCACGGAGA CGGAACCTGC TCTGAAATAC 600  
CTTCTGCACA GAGGGTTGCC TGGAGGCAGA AATGCTTCTG TGCCCCAGAT CCTCATCATC 660  
GTCAGTATG GGAAGTCCCA GGGGGATGTG GCACTGCCAT CCAAGCAGCT GAAGGAAAGG 720  
GGTGTCACTG TGTGTGCTGT GGGGGTCAGG TTTCCAGGT GGGAGGAGCT GCATGCACTG 780  
GCCAGCGAGC CTAGAGGGCA GCACGTGCTG TTGGCTGAGC AGGTGGAGGA TGCCACCAAC 840  
GGCCTCTTCA GCACCCTCAG CAGCTCGGCC ATCTGCTCCA GCGCCACGCC AGCTGGGAGC 900  
CCCAGACTTG TCTTCATGGA GCGGTTAATG GGCATCTCTC TGATAGGCCC CTGTGACTCG 960  
CAGCCCTGCC AGAATGGAGG CACATGTGTT CCAGAAGGAC TGGACGGCTA CCAGTGCCTC 1020  
TGCCCCCTGG CCTTTGGAGG GGAGGCTAAC TGTGCCCTGA AGCTGAGCCT GGAATGCAGG 1080  
GTCGACCTCC TCTTCTGTCT GGACAGCTCT GCGGGCACC A CTCTGGACGG CTCTCTGCGG 1140  
GCCAAGTCT TCGTGAAGCG GTTTGTGCGG CCGGTGCTGA GCGAGGACTC TCGGGCCCCG 1200  
GTGGGTGTGG CCACATACAG CAGGGAGCTG CTGGTGGCGG TGCCTGTGGG GGAGTACCAG 1260  
GATGTGCCTG ACCTGGTCTG GAGCCTCGAT GGCATTCCCT TCCGTGGTGG CCCACCCTG 1320  
ACGGGCAGTG CCTTGCAGCA GCGGCGAGAG CGTGGCTTCG GGAGCGCCAC CAGGACAGGC 1380  
CAGGACCGGC CACGTAGAGT GGTGGTTTTG CTTACTGAGT CACACTCCGA GGATGAGGTT 1440  
GCGGGCCAG CGCGTCACGC AAGGGCGCGA GAGCTGCTCC TGCTGGGTGT AGGCAGTGAG 1500  
GCCGTGCGGG CAGAGCTGGA GGAGATCACA GGCAGCCCAA AGCATGTGAT GGTCTACTCG 1560  
GATCCTCAGG ATCTGTTCAA CCAAATCCCT GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG 1620  
CGGCCAGGGT GCCGGACACA AGCCCTGGAC CTCGTCTTCA TGTGAGACAC CTCTGCCTCA 1680  
GTAGGGCCCC AGAATTTTGC TCAGATGCAG AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT 1740  
GAGGTGAACC CTGACGTGAC ACAGGTCCGC CTGGTGGTGT ATGGCAGCCA GGTGCAGACT 1800  
GCCTTCGGGG TGGACACCAA ACCACCCGG GCTGCGATGC TGCGGGCCAT TAGCCAGGCC 1860  
CCCTACCTAG GTGGGGTGGG CTAGGCCGGC ACCGCCCTGC TGACATCTA TGACAAAGTG 1920  
ATGACCGTCC AGAGGGGTGC CCGGCCTGGT GTCCCCAAAG CTGTGGTGGT GCTCACAGGC 1980  
GGGAGAGGCG CAGAGGATGC AGCCGTTCCT GCCCAGAAGC TGAGGAACAA TGGCATCTCT 2040  
GTCTTGGTCTG TGGGCGTGGG GCCTGTCTTA AGTGAGGGTC TGCGGAGGCT TGCAAGTCCC 2100  
CGGATTTCCC TGATCCACGT GGCAGCTTAC GCCGACCTGC GGTACCACCA GGACGTGCTC 2160  
ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG CCAGTCAACC TCTGCAAACC CAGCCCGTGC 2220  
ATGAATGAGG GCAGCTGCGT CTAAGCAAGT GGGAGCTACC GCTGCAAGTG TCGGATGGC 2280  
TGGAGAGGCC CCCACTGCGA AACCCGTGAG TGAGGCTCTT GCTCTGTATG TGTGAGCCAG 2340  
GGATGGATTG TTGAGACGCC CCTGAGGCAC ATGGCTCCCG TGCAGGAGGG CAGCAGCCGT 2400  
ACCCCTCCCA GCAACTACAG AGAAGGCCTG GGCAGTAAA TGGTGCCTAC CTCTCTGGAAT 2460  
GTCTGTGCCC CAGGTCCTTA G

Seq ID NO: 441 Protein sequence  
Protein Accession #: XP\_061091.1

1 11 21 31 41 51  
| | | | |  
MPNTSGTTRI EIWLLQEPPG HRALVAALLP VSPSPALALA PGYPPVPAAD DRFTLPMIGG 60  
QMHGEKVDLW SLGVLCEYFL VGKPPFEANE VHSVKEITIG ISAASKMMWC SAAVDIMFLL 120  
DGSNSVFKGS PERSKHFAIT VCDGLDISPE RVRVGAQFQS STPHLEFPLD SFSTQQEVKA 180  
RIKRMVFKGG RTETELALKY LLHRGLPGGR NASVPQILII VTDGKSQGDV ALPSKQLKER 240  
GVTVFAVGVR FPRWEELHAL ASEPRGQHVLA LAEQVEDATN GLFSTLSSSA ICSSATPAGS 300  
PELVFMERLM GISLIGPCDS QPCQNGGTCV PEGLDGYQCL CPLAFGGGAN CALKLSLECR 360  
VDLLFLLDSS AGTTLDGFLR AKVFKRFRV AVLSSESRAR VGVATYSREL LVAVPVGEYQ 420  
DVPDLVWSD GIPFRGGPTL TGSALRQAAE RGFGSATRTG QDRPRRVVVL LTESHSEDEV 480  
AGPARHARAR ELLLLGVGSE AVRALEEIT GSPKHVMVYS DPQDLFNQIP ELQKLCRSRQ 540  
RPGCRTQALD LVFMLDTSAS VGPENFAQMQ SFVRSCALQF EVNPDVTQVG LVVYGSQVQT 600  
AFGLDTKPTR AAMLRAISQA PYLGGVGSAG TALLHIYDKV MTVQRGARPG VPKAVVLTG 660  
GRGAEDAAPV AQKLRNNGIS VLVVGVGPVL SEGLRRLAGP RDSLIHVAAY ADLRYHQDVL 720



IEWLCEAKQ PVNLCKPSPC MNEGSCVLQN GSYRCKCRDG WEGPHCENRE WSSCSVCVSQ 780  
GWILETPLRH MAPVQEGSSR TPPSNYREGL GTEMVPTFWN VCAPGP

Seq ID NO: 442 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..2424

10 1 11 21 31 41 51  
| | | | | |  
ATGCCCCCTT TCCTGTTGCT GGAGGCCGTC TGTGTTTTTC TGTTTTCCAG AGTGCCCCCA 60  
TCTCTCCCTC TCCAGGAAGT CCATGTAAGC AAAGAAACCA TCGGGAAGAT TTCAGCTGCC 120  
AGCAAAATGA TGTGGTGCTC GGCTGCAGTG GACATCATGT TTCTGTTAGA TGGGTCTAAC 180  
AGCGTCGGGA AAGGGAGCTT TGAAAGGTCC AAGCACTTTC CCATCACAGT CTGTGACGGT 240  
CTGGACATCA GCCCCGAGAG GGTGAGAGTG GGAGCATTTC AGTTCAGTTC CACTCCTCAT 300  
15 CTGGAATTCC CTTGGATTTC ATTTTCAACC CAACAGGAAG TGAAGGCAAG AATCAAGAGG 360  
ATGGTTTTCA AAGGAGGGCG CACGGAGACG GAACTTGCTC TGAATAACCT TCTGCACAGA 420  
GGGTTCCTG GAGGCAGAAA TGCTTCTGTG CCCCAGATCC TCATCATCGT CACTGATGGG 480  
AAGTCCCAGG GGGATGTGGC ACTGCCATCC AAGCAGCTGA AGGAAAGGGG TGTCACTGTG 540  
TTTGCTGTGG GGGTCAGGTT TCCCAGGTGG GAGGAGCTGC ATGCACTGGC CAGCGAGCCT 600  
20 AGAGGGCAGC ACGTGCTGTT GGCTGAGCAG GTGGAGGATG CCACCAACGG CCTCTTCAGC 660  
ACCTCAGCA GCTCGCCAT CTGCTCCAGC GCCACGCCAG ACTGCAGGGT CGAGGCTCAC 720  
CCCTGTGAGC ACAGGACGCT GGAGATGGTC CGGGAGTTTC CTGGCAATGC CCCATGCTGG 780  
AGAGGATCGC GCGGACCCCT TGCGGTGCTG GCTGCACACT GTCCCTTCTA CAGCTGGAAG 840  
AGAGTGTTC TAACCCACCC TGCCACCTGC TACAGGACCA CCTGCCCAGG CCCCTGTGAC 900  
25 TCGCAGCCCT GCCAGAATGG AGGCACATGT GTTCCAGAAG GACTGGACGG CTACCAAGTGC 960  
CTCTGCCCGC TGGCCTTTGG AGGGGAGGCT AACTGTGCCC TGAAGCTGAG CCTGGAATGC 1020  
AGGGTCGACC TCCTCTTCTT GCTGGACAGC TCTGCGGGCA CCACTCTGGA CGGCTTCCTG 1080  
CGGGCCAAAG TCTTCGTGAA GCGGTTTGTG CCGGCCGTGC TGAGCGAGGA CTCTCGGGCC 1140  
30 CGAGTGGGTG TGGCCACATA CAGCAGGGAG CTGCTGGTGG CCGTGCTGTG GGGGGAGTAC 1200  
CAGGATGTGC CTGACCTGGT CTGGAGCCTC GATGGCATTC CCTTCCGTGG TGGCCCCACC 1260  
CTGACGGGCA GTGCCTTGCG GCAGGCGGCA GAGCGTGGCT TCGGGAGCGC CACCAGGACA 1320  
GGCCAGGACC GGCCACGTAG AGTGGTGGTT TTGCTCACTG AGTCACACTC CGAGGATGAG 1380  
GTTGCGGGCC CAGCGCGTCA CGCAAGGGCG CGAGAGCTGC TCCTGCTGGG TGTAGGCAGT 1440  
GAGGCCGTGC GGGCAGAGCT GGAGCAGGCC CAAAGCATGT GATGGTCTAC 1500  
35 TCGGATCCTC AGGATCTGTT CAACCAAAATC CCTGAGCTGC AGGGGAAGCT GTGCAGCCGG 1560  
CAGCGGCCAG GGTGCCGGAC ACAAGCCCTG GACCTCGTCT TCATGTTGGA CACCTCTGCC 1620  
TCAGTAGGGC CCGAGAATTG TGCTCAGATG CAGAGCTTTG TGAGAAGCTG TGCCCTCCAG 1680  
TTTGAGGTGA ACCCTGACGT GACACAGGTC GGCCTGGTGG TGTATGGCAG CCAGGTGCAG 1740  
ACTGCCTTCG GGCTGGACAC CAAACCCACC CGGGCTGCGA TGCTGCGGGC CATTAGCCAG 1800  
40 GCCCCCTACC TAGGTGGGGT GGGCTCAGCC GGCACGCCCC TGCTGCACAT CTATGACAAA 1860  
GTGATGACCG TCCAGAGGGG TGCCCGGCCT GGTGTCCCCA AAGCTGTGGT GGTGCTCACA 1920  
GGCGGGAGAG GCGCAGAGGA TGCAGCCGTT CCTGCCAGA AGCTGAGGAA CAATGGCATC 1980  
TCTGTCTTGG TCGTGGGCGT GGGGCTGTC CTAAGTGAGG GTCTGCGGAG GCTTGCAGGT 2040  
CCCCGGGATT CCCTGATCCA CGTGGCAGCT TACGCCGACC TGCGGTACCA CCAGGACGTG 2100  
45 CTCATTGAGT GGCTGTGTGG AGAAGCCAAG CAGCCAGTCA ACCTCTGCAA ACCCAGCCCG 2160  
TGCATGAATG AGGCAGCTG CGTCTGAGC AATGGGAGCT ACCGCTGCAA GTGTGCGGAT 2220  
GGCTGGGAGG GCGCCCACTG CGAGAACCCT GAGTGGAGCT CTTGCTCTGT ATGTGTGAGC 2280  
CAGGGATGGA TTCTTGAGAC GCGCCCTGAG CACATGGCTC CCGTGCAGGA GGGCAGCAGC 2340  
50 CGTACCCCTC CCAGCAACTA CAGAGAAGGC CTGGGCACTG AAATGGTGCC TACCTTCTGG 2400  
AATGTCTGTG CCCAGGTCC TTAG

Seq ID NO: 443 Protein sequence  
Protein Accession #: Eos sequence

55 1 11 21 31 41 51  
| | | | | |  
MPPFLLLEAV CVFLFSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN 60  
SVGKGSFERS KHFAITVCDG LDISPERVRV GAFQFSSTPH LEPLDSFST QQEVKARIKR 120  
60 MVFKGRTET ELALKYLLHR GLPGGRNASV PQILIIIVTDG KSQGDVALPS KQLKERVTV 180  
FVGVRFPRW EELHALASEP RGQHVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH 240  
PCEHRTLEMV REFAGNAPCW RGSRRTLAVL AAHCPFYSWK RVFLTHPATC YRTTCPGPCD 300  
SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEC RVDLLFLDLS SAGTTLDGFL 360  
RAKVFKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPDLVWSL DGIPFRGGPT 420  
65 LTGSALRQAA ERFGSATRT GQDRPRRVVV LLTESHSEDE VAGPARHARA RELLLLGVGS 480  
EAVRAELEEI TGSPKHVMVY SDPQDLFNQI PELQKLCRSL QRPGRCTQAL DLVFMLD TSA 540  
SVGPENFAQM QSFVRSCALQ FEVNPDTVQI GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ 600  
APYLGGVGS A GTALLHIYDK VMTVQRGARP GVPKAVVVL T GGRGAEDAAV PAQKLNRNGI 660  
SVLVVGVGPV LSEGLRRLAG PRDSLHIVAA YADLRYHQDV LIEWLCEAK QPVNLCKPSP 720  
70 CMNEGSCVLQ NGSYRCKCRD GWEGPHCENR EWSSCSVCVS QGWILETPLR HMAPVQEGSS 780  
RTPPSNYREG LGTEMVPTFW NVCAPGP

Seq ID NO: 444 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 89..2356

75 1 11 21 31 41 51  
| | | | | |  
GCCCCCTGGC CCGAGCCGCG CCGGGTCTG TGAGTAGAGC CGCCCGGGCA CCGAGCGCTG 60  
GTGCGCGCTC TCCTTCCGTT ATATCAACAT GCCCCCTTTC CTGTTGCTGG AAGCCGTCTG 120  
80 TGTTCCTCTG TTTTCCAGAG TGCCCCCATC TCTCCCTCTC CAGGAAGTCC ATGTAAGCAA 180  
AGAAACCATC GGAAGATTG CAGCTGCCAG CAAAATGATG TGGTGCTCGG CTGCAGTGGA 240  
CATCATGTTT CTGTTAGATG GGTCTAACAG CGTCGGGAAA GGGAGCTTTG AAAGGTCCAA 300  
GCACTTTGCC ATCAGAGTCT GTGACGGTCT GGACATCAGC CCCGAGAGGG TCAGAGTGGG 360  
85 AGCATTCAG TTCAGTTCCA CTCTCATCT GGAATCCCC TTGGATTCTT TTTCAACCCA 420  
ACAGGAAGTG AAGGCAAGAA TCAAGAGGAT GGTTCCTCAA GGAGGGCGCA CCGAGACGGA 480  
ACTTGCTCTG AAATACCTTC TGACAGAGAG GTTGCTGGA GGCAGAAATG CTTCTGTGCC 540  
CCAGATCCTC ATCATCGTCA CTGATGGGAA GTCCCAGGGG GATGTGGCAC TGCCATCCAA 600



	GCAGCTGAAG	GAAAGGGGTG	TCACTGTGTT	TGCTGTGGGG	GTCAGGTTTC	CCAGGTGGGA	660
	GGAGCTGCAT	GCACTGGCCA	GCGAGCCTAG	AGGGCAGCAC	GTGCTGTTGG	CTGAGCAGGT	720
	GGAGGATGCC	ACCAACGGCC	TCTTCAGCAC	CCTCAGCAGC	TCGGCCATCT	GCTCCAGCGC	780
5	CACGCCAGAC	TGCAGGGTCG	AGGCTCACCC	CTGTGAGCAC	AGGACGCTGG	AGATGGTCCG	840
	GGAGTTGCT	GGCAATGCCC	CATGCTGGAG	AGGATCGCGG	CGGACCCTTG	CGGTGCTGGC	900
	TGCACACTGT	CCCTTCTACA	GCTGGAAGAG	AGTGTTCCTA	ACCCACCCTG	CCACCTGCTA	960
	CAGGACCACC	TGCCCAGGCC	CCTGTGACTC	GCAGCCCTGC	CAGAATGGAG	GCACATGTGT	1020
	TCCAGAAGGA	CTGGACGGCT	ACCAGTGCCT	CTGCCCCTG	GCCTTTGGAG	GGGAGGCTAA	1080
10	CTGTGCCCTG	AAGCTGAGCC	TGGAATGCAG	GGTCGACCTC	CTCTTCTGTC	TGGACAGCTC	1140
	TGCGGGCACC	ACTCTGGACG	GCTTCTGCG	GGCCAAAGTC	TTCGTGAAGC	GGTTTGTGCG	1200
	GGCCGTGCTG	AGCGAGGACT	CTCGGGCCCG	AGTGGGTGTG	GCCACATACA	GCAGGGAGCT	1260
	GCTGGTGGCG	GTGCCTGTGG	GGGAGTACCA	GGATGTGCCT	GACCTGGTCT	GGAGCCTCGA	1320
	TGGCATTCCC	TTCCGTGGTG	GCCCCACCCT	GACGGGACGT	GCCTTGCGGC	AGGCGGCAGA	1380
	GCGTGGCTTC	GGGAGCGCCA	CCAGGACAGG	CCAGGACCGG	CCACGTAGAG	TGGTGGTTTT	1440
15	GCTCACTGAG	TCACACTCCG	AGGATGAGGT	TGCGGGCCCA	GCGCGTCACG	CAAGGGCGCG	1500
	AGAGCTGCTC	CTGCTGGGTG	TAGGCAGTGA	GGCCGTGCGG	GCAGAGCTGG	AGGAGATCAC	1560
	AGGCAGCCCA	AAGCATGTGA	TGGTCTACTC	GGATCCTCAG	GATCTGTTCA	ACCAAATCCC	1620
	TGAGCTGCAG	GGGAAGCTGT	GCAGCCGGCA	GCGGCCAGGG	TGCCCGACAC	AAGCCCTGGA	1680
20	CCTCGTCTTC	ATGTTGGACA	CCTCTGCCTC	AGTAGGGCCC	GAGAATTTTG	CTCAGATGCA	1740
	GAGCTTTGTG	AGAAGCTGTG	CCCTCCAGTT	TGAGGTGAAC	CCTGACGTGA	CACAGGTCGG	1800
	CCTGTTGGTG	TATGGCAGCC	AGGTGCAGCC	TGCCTTCGGG	CTGGACACCA	AACCCACCCG	1860
	GGCTGCGATG	CTGCGGGCCA	TTAGCCAGGC	CCCCTACCTA	GGTGGGGTGG	GCTCAGCCGG	1920
	CACCGCCCTG	CTGCACATCT	ATGACAAAGT	GATGACCGTC	CAGAGGGGTG	CCCGGCCTGG	1980
25	TGTCCCCAAA	GCTGTGGTGG	TGCTCACAGG	CGGGAGAGGC	GCAGAGGATG	CAGCCGTTCC	2040
	TGCCCAGAAG	CTGAGGAACA	ATGGCATCTC	TGTCTTGGTC	GTGGGCGTGG	GGCCTGTCTT	2100
	AAGTGAGGGT	CTGCGGAGGC	TTGCAGGTCC	CCGGGATTCC	CTGATCCACG	TGGCAGCTTA	2160
	CGCCGACCTG	CGGTACCAAC	AGGACGTGCT	CATTGAGTGG	CTGTGTGGAG	AAGCCAAGCA	2220
	GCCAGTCAAC	CTCTGCAAAAC	CCAGCCCGTG	CATGAATGAG	GGCAGCTGCG	TCCTGCAGAA	2280
30	TGGGAGCTAC	CGCTGCAAGT	GTCGGGATGG	CTGGGAGGGC	CCCCACTGCG	AGAACCATT	2340
	CTTGAGACGC	CCCTGAGGCA	CATGGCTCCC	GTGCAGGAGG	GCAGAGCCG	TACCCCTCCC	2400
	AGCAACTACA	GAGAAGGCCT	GGGCACTGAA	ATGGTGCCTA	CCTTCTGGAA	TGTCTGTGCC	2460
	CCAGGTCCTT	AGAATGTCTG	CTTCCCGCCG	TGGCCAGGAC	CACTATTCTC	ACTGAGGGAG	2520
	GAGGATGTCC	CAACTGCAGC	CATGCTGCTT	AGAGCAAGA	AAGCAGCTGA	TGTCAACCAC	2580
35	AAACGATGTT	GTTGAAAAGT	TTTGATGTGT	AAGTAAATAC	CCACTTTCTG	TACCTGCTGT	2640
	GCCTTGTTGA	GGCTATGTCA	TCTGCCACCT	TTCCCTTGAG	GATAAACAAAG	GGGTCTTGAA	2700
	GACTTAAATT	TAGCGGCCTG	ACGTTCCCTT	GCACACAATC	AATGCTCGCC	AGAATGTTGT	2760
	TGACACAGTA	ATGCCAGCA	GAGGCCCTTA	CTAGAGCATC	CTTTGGACGG		

Seq ID NO: 445 Protein sequence  
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
45	MPPFLLLEAV	CVFLFSRVPP	SLPLQEVHVS	KETIGKISAA	SKMMWCSAAV	DIMFLLDGSN	60
	SVGKGSFERS	KHFAITVCDG	LDISPERNRV	GAQFSSSTPH	LEFPLDSFST	QDEVKARIKR	120
	MVFKGGRTE	ELALKYLLHR	GLPGRNASV	PQILIIIVTDG	KSQGDVALPS	KQLKERGVTV	180
	FAVGVRFP	EELHALASEP	RQHVLLAEQ	VEDATNGLFS	TLSSSAICSS	ATPDCRVEAH	240
	PCEHRTLE	REFAGNAPCW	RGSRRTLAVL	AAHCPFYSWK	RVFLTHPATC	YRTTCGP	300
50	SQPCQNGG	VPEGLDGYQC	LCPLAFGGEA	NCALKLSLEC	RVDLLFLLDS	SAGTTLDGFL	360
	RAKVFKRFV	RAVLSEDSRA	RVGVATYSRE	LLVAVPVGEY	QDVPDLVWSL	DGIPFRGGPT	420
	LTGSALRQAA	ERGFSA	GQDRPRRVV	LLTESHSEDE	VAGPARHARA	RELLLLGVGS	480
	EAVRAELEEI	TGSPKHVMVY	SDPQDLFNQI	PELQGLKCSR	QRPGRCTQAL	DLVFM	540
	SVGPENFAQM	QSFVRSCALQ	FEVNP	GLVVG	TAFGLDTKPT	RAAMLRAISQ	600
55	APYLGGV	GTALLHIYDK	VMTVQ	GVPKAVV	GGRGAEDAAV	PAQKL	660
	SVLVVGVGPV	LSEGLRRLAG	PRDSLIHVAA	YADLRYHQDV	LIEWLCGEAK	QPVNLCKPSP	720
	CMNEGSCVLQ	NGSYRCKCRD	GWEGPHCENR	FLRRP			

Seq ID NO: 446 DNA sequence  
Nucleic Acid Accession #: NM\_031942.1  
Coding sequence: 145..1260

	1	11	21	31	41	51	
65	CCCAGCCCC	GCCCCCTCCG	GCCCGGGTCG	GCGCGCCAG	CCTGCCAGCC	GCGCTGCTGC	60
	TGCTCCTCCT	GCTGTGGGAC	GCGTGACCGC	GCGGCTGCTC	CGCTCTCCCC	GCTCCAAGCG	120
	CCGATCTGGG	CACCCGCCAC	CAGCATGGAC	GCTCGCCGCG	TGCCCGAGAA	AGATCTCAGA	180
	GTAAAGAAGA	ACTTAAAGAA	ATTCAGATAT	GTGAAGTTGA	TTTCCATGGA	AACCTCGTCA	240
	TCCTCTGATG	ACAGTTGTGA	CAGCTTTGCT	TCTGATAATT	TTGCAACAC	GAGGCTGCAG	300
70	TCAGTTCCGG	AAGGCTGTAG	GACCCGCAGC	CAGTGCAGGC	ACTCTGGACC	TCTCAGGGTG	360
	GCGATGAAGT	TTCCAGCGCG	GAGTACCAGG	GGAGCAACCA	ACAAAAAAGC	AGAGTCCCGC	420
	CAGCCCTCAG	AGAATTCTGT	GACTGATTCC	AACTCCGATT	CAGAAGATGA	AAGTGGAATG	480
	AATTTTTTGG	AGAAAAGGGC	TTTAAATATA	AAGCAAAACA	AAGCAATGCT	TGCAAACTC	540
	ATGTCTGAAT	TAGAAAGCTT	CCCTGGCTCG	TTCCGTGGAA	GACATCCCTT	CCCAGGCTCC	600
75	GACTCACAA	CAAGGAGACC	GCGAAGGCGT	ACATTCCCGG	GTGTTGCTTC	CAGGAGAAAC	660
	CCTGAACGGA	GAGCTCGTCC	TCTTACCAGG	TCAAGGTCCC	GGATCCTCGG	GTCCCTTGAC	720
	GCTCTACCCA	TGGAGGAGGA	GGAGGAAGAG	GATAAGTACA	TGTTGGT	AAAGAGGAAG	780
	ACCGTGGATG	GCTACATGAA	TGAAGATGAC	CTGCCAGAA	GCCGTCGCTC	CAGATCATCC	840
	GTGACCCTTC	CGCATATAAT	TCGCCCAGTG	GAAGAAATTA	CAGAGGAGGA	GTTGGAGAAC	900
80	GTCTGCAGCA	ATTCTCGAGA	GAAGATATAT	AACCGTTCAC	TGGGCTCTAC	TTGTCATCAA	960
	TGCCGTCAGA	AGACTATTGA	TACCAAAACA	AACTGCAGAA	ACCCAGACTG	CTGGGGCGTT	1020
	CGAGGCCAGT	TCTGTGGCCC	CTGCCTTCGA	AACCGTTATG	GTGAAGAGGT	CAGGGATGCT	1080
	CTGCTGGATC	CGAACTGGCA	TTGCCCGCCT	TGTCGAGGAA	TCTGCAACTG	CAGTTTCTGC	1140
	CGGCACGAG	ATGGACGGTG	TGCGACTGGT	GTCTTGTGT	ATTTAGCCAA	ATATCATGGC	1200
85	TTTGGGAATG	TCATGCCTA	CTGAAAAGC	GTGAAACAGG	AATTTGAAAT	GCAAGCATAA	1260
	TATCTGGA	ATTGCTGCC	TGCCTTCTAC	TTCTCAAATC	TTTCTGTGAA	AAGTTTCCAA	1320
	TTTTTCACT	GAAACCTGAG	TTAAAAATCT	TGATGATCAG	CCTGTTTCAT	AAGAACTCC	1380
	AATCAAGTTA	ATCTTAGCAG	ACATGTGTTT	CTGGAGCATC	ACAGAAGGTA	TATTGCTAGT	1440



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60

TACACTTTGC CCTCCTGCAG TTTCTTCTCT GCTCCCAACC CCCATCTCAT AGCATCCCCC 1500  
TCTATTTCCA ATGCTCCTCT CCAACCGCTT AGTTTCTGAA TTTCTTTTAA ATTACAGTTT 1560  
TATGAAAGCA TATTTTATTT ACTTGGTGTT GAAATAGCCC TCATAAAACC TAAGCACTTG 1620  
GAAACACAAT AATAGTATTA ACTAACTAGA TCTATTGAAT TTCAGAGAAG AGCCTTCTAA 1680  
CTTGTTTACA CAAAAACGAG TATGATTTAG CACTCATACT AGTTGAAATT TTTAATAGAA 1740  
TCAAGGCACA AAAGTCTTAA AACCATGTGG AAAAATTAGG TAATTATTGC AGATTGATGT 1800  
CTCTCAATCC CATGTATTGC GCTTATGTTA CAAGTTGTTG TCACAGTTGA GACTTAATTT 1860  
CTCCTAATTT CTTCTGCCCC AAGGGTAAGT GSDSCGTCCA GCTTACACGA TCATAATTCA 1920  
AAGGTTGGTG GGCAATGTAA TACTTAATTA AAATAATGAT GGAAGAGCTA TCTGGAGATT 1980  
ATGAGTAAGC TGATTTGAAT TTTCAGTATA AAACCTTAGT ATAATTGTAG TTTGCAAAGT 2040  
TTATTTTCAGT TCACATGTAA GGTATTGCAA ATAAATTCTT GGACAATTTT GTATGGAAAC 2100  
TTGATATTAA AAAC TAGTCT GTGGTCTTT GCAGTTTCTT GTAAATTTAT AAACCAGGCA 2160  
CAAGGTTCAA GTTTAGATTT TAAGCACTTT TATAACAATG ATAAGTGCCT TTTTGGAGAT 2220  
GTAACCTTTA GCAGTTTGTG AACCTGACAT CTCTGCCAGT CTAGTTTCTG GGCAGGTTTC 2280  
CTGTGTCAGT ATTTCCCTCT CTTTTCGAT TAATCAAGGT ATTTGGTAGA GGTGGAATCT 2340  
AAGTGTGTTG ATGTCCAATT TACTGTCATA TGTAACCAT TGCTGTGCCA TTCAATGTTT 2400  
GATGCATAAT TGGACCTTGA ATCGATAAGT GTAAATACAG CTTTGTATCT GTAATGCTTT 2460  
TATACAAAAG TTTATTTTAA TAATAAAATG TTTGTTCTAA AAAAAAAAAA

Seq ID NO: 447 Protein sequence  
Protein Accession #: NP\_114148.1

1 11 21 31 41 51  
| | | | | |  
MDARRVPQKD LRVKKNLKKF RYVKLISMET SSSSDDSCDS FASDNFANTR LQSVREGCRT 60  
RSQCRHSGPL RVAMKFPARS TRGATNKKAE SRQPSSENSVT DSNSDSEDES GMNFLEKRAL 120  
NIKQNKAMLA KLMSELESFP GSFRGRHPLP GSDSQSRRPR RRTFPGVASR RNPERRARPL 180  
TRSRSRILGS LDALPMEEEE EEDKYMVLVR RKTVDGYMNE DDLPRSRRSR SSVTLPHIIR 240  
PVEEITEEEL ENVCSNSREK IYNRSLGSTC HQCRQKTIDT KTNCRNPDCW GVRGQFCGPC 300  
LRNRYGEEVR DALLDPNWHC PPCRGICNCS FCRQRDGRCA TGVLVYLAKY HGFGNVHAYL 360  
KSLKQEFEMQ A

Seq ID NO: 448 DNA sequence  
Nucleic Acid Accession #: NM\_019894  
Coding sequence: 1..1314

1 11 21 31 41 51  
| | | | | |  
ATGTTACAGG ATCCTGACAG TGATCAACCT CTGAACAGCC TCGATGTCAA ACCCCTGCGC 60  
AAACCCCGTA TCCCCTGGA GACCTTCAGA AAGGTGGGGA TCCCCATCAT CATAGCACTA 120  
CTGAGCCTGG CGAGTATCAT CATTGTGGTT GTCTCATCA AGGTGATTCT GGATAAATAC 180  
TACTTCTCT GCGGGCAGCC TCTCCACTTC ATCCCGAGGA AGCAGCTGTG TGACGGAGAG 240  
CTGGACTGTC CTTGGGGGGA GGACGAGGAG CACTGTGTCA AGAGCTTCCC CGAAGGGCCT 300  
GCAGTGGCAG TCCGCCTCTC CAAGGACCGA TCCACACTGC AGGTGCTGGA CTCGGCCACA 360  
GGGAAGTGGT TCTCTGCCTG TTTCGACAAC TTCACAGAAG CTCTCGCTGA GACAGCCTGT 420  
AGGCAGATGG GCTACAGCAG CAAACCCACT TTCAGAGCTG TGGAGATTGG CCCAGACCAG 480  
GATCTGGATG TTGTTGAAAT CACAGAAAAC AGCCAGGAGC TTCGCATGCG GAACTCAAGT 540  
GGGCCCTGTC TCTCAGGCTC CTTGGTCTCC CTGCACTGTC TTGCCTGTGG GAAGAGCCTG 600  
AAGACCCCC GTGTGGTGGG TGGGGAGGAG GCCTCTGTGG ATTCTTGGCC TTGGCAGGTC 660  
AGCATCCAGT ACGACAAACA GCACGTCTGT GGAGGGAGCA TCCTGGACCC CCACTGGGTC 720  
CTCAGGCGAG CCCACTGCTT CAGGAAACAT ACCGATGTGT TCAACTGGAA GGTGCGGGCA 780  
GGCTCAGACA AACTGGGCAG CTTCCCATCC CTGGCTGTGG CCAAGATCAT CATCATTGAA 840  
TTCAACCCCA TGTACCCCAA AGACAAATGAC ATCGCCCTCA TGAAGCTGCA GTTCCCACTC 900  
ACTTTCTCAG GCACAGTCAG GCCCATCTGT CTGCCCTTCT TTGATGAGGA GCTCACTCCA 960  
GCCACCCAC TCTGGATCAT TGGATGGGG TTTACGAAGC AGAATGGAGG GAAGATGTCT 1020  
GACATACTGC TGCAGGCGTC AGTCCAGGTC ATTGACAGCA CACGGTGCAA TGCAGACGAT 1080  
GCGTACCAGG GGGAAAGTCAC CGAGAAGATG ATGTGTGCAG GCATCCCGGA AGGGGGTGTG 1140  
GACACCTGCC AGGGTGACAG TGGTGGGCCC CTGATGTACC AATCTGACCA GTGGCATGTG 1200  
GTGGGCATCG TTAGCTGGGG CTATGGCTGC GGGGCCCCGA GCACCCAGG AGTATACACC 1260  
AAGGTCTCAG CCTATCTCAA CTGGATCTAC AATGTCTGGA AGGCTGAGCT GTAA

Seq ID NO: 449 Protein sequence  
Protein Accession #: NP\_063947.1

1 11 21 31 41 51  
| | | | | |  
MLQDPDSQDQ LNSLDVKPLR KPRIPMETFR KVGIPIIIAL LSLASIIIV VLIKVILDKY 60  
YFLCGQPLHF IPRKQLCDGE LDCPLGEDEE HCVKSFPEGP AVAVRLSKDR STLQVLD SAT 120  
GNWFSACFDN FTEALAETAC RQMGYSKPT FRAVEIGPDQ DLDVVEITEN SQELMRNSS 180  
GPCLSGSLVS LHCLACGKSL KTPRVVGEE ASVDSWPWQV SIQYDKQHVC GGSILDPHWV 240  
LTAHCFRKH TDVFNWKVRA GSDKLGSPS LAVAKIIIE FNPMPKDND IALMKLQFPL 300  
TFSGTVRPIC LPFFDEELTP ATPLWIIWGW FTKQNGGKMS DILLQASVQV IDSTRCNADD 360  
AYQGEVTEKM MCAGIPEGGV DTCQGDGGP LMYQSDQWHV VGIVSWG YGC GGPSTPGVYT 420

Seq ID NO: 450 DNA sequence  
Nucleic Acid Accession #: XM\_051860.2  
Coding sequence: 52..3042

1 11 21 31 41 51  
| | | | | |  
GCTCACCCAG GAAAAATATG CAATCGTCCC ATTGATATAC AGGCCACTAC AATGGATGGA 60  
GTTAACCTCA GCACCGAGGT TGTCTACAAA AAAGGCCAGG ATTATAGGTT TGCTTGCTAC 120  
GACCGGGGCA GAGCCTGCCG GAGCTACCGT GTACGTTCC TCTGTGGGAA GCCTGTGAGG 180  
CCCAACTCA CAGTCAACAT TGACACCAAT GTGAACAGCA CCATTCTGAA CTTGGAGGAT 240  
AATGTACAGT CATGGAAACC TGGAGATACC CTGGTCATTG CCAGTACTGA TTACTCCATG 300  
TACCAGGCAG AAGAGTTCCA GGTGCTTCCC TGCAGATCCT GCGCCCCCAA CCAGGTCAAA 360



	GTGGCAGGGA	AACCAATGTA	CCTGCACATC	GGGGAGGAGA	TAGACGGCGT	GGACATGCGG	420
	GCGGAGGTTG	GGCTTCTGAG	CCGGAACATC	ATAGTGTATG	GGGAGATGGA	GGACAAATGC	480
	TACCCCTACA	GAAACCACAT	CTGCAATTTC	TTTGACTTCG	ATACCTTTGG	GGGCCACATC	540
5	AAGTTTGCTC	TGGGATTTAA	GGCAGCACAC	TTGGAGGGCA	CGGAGCTGAA	GCATATGGGA	600
	CAGCAGCTGG	TGGGTCACTA	CCCATTTCAC	TTCCACCTGG	CCGGTGATGT	AGACGAAAGG	660
	GGAGGTTATG	ACCCACCCAC	ATACATCAGG	GACCTCTCCA	TCCATCATA	ATTCTCTCGC	720
	TGCGTCACAG	TCCATGGCTC	CAATGGCTTG	TTGATCAAGG	ACGTTGTGGG	CTATAACTCT	780
	TTGGGCCACT	GCTTCTTCAC	GGAAGATGGG	CCGGAGGAAC	GCAACACTTT	TGACCACTGT	840
10	CTTGCCCTCC	TTGTCAAGTC	TGGAACCCCTC	CTCCCCCTCG	ACCGTGACAG	CAAGATGTGC	900
	AAGATGATCA	CAGGAGACTC	CTACCCAGGG	TACATCCCCA	AGCCCAGGCA	AGACTGCAAT	960
	GCTGTGTCCA	CCTTCTGGAT	GGCCAATCCC	AACAACAACC	TCATCAACTG	TGCCGCTGCA	1020
	GGATCTGAGG	AAACTGGATT	TTGGTTTAT	TTTCACCACG	TACCAACGGG	CCCTCCCGTG	1080
	GGAAATGTACT	CCCCAGGTTA	TTTCAGAGCAC	ATTCCACTGG	GAAAATTCTA	TAACAACCGA	1140
	GCACATTCCA	ACTACCGGGC	TGGCATGATC	ATAGACAACG	GAGTCAAAAC	CACCGAGGCC	1200
15	TCTGCCAAGG	ACAAGCGGCC	GTTCCCTCTCA	ATCATCTCTG	CCAGATACAG	CCCTCACCAG	1260
	GACGCCGACC	CGCTGAAGCC	CCGGGAGCCG	GCCATCATCA	GACACTTCAT	TGCCTACAAG	1320
	AACCAGGACC	ACGGGGCCCTG	GCTGCGCGGC	GGGGATGTGT	GGCTGGACAG	CTGCCGTTT	1380
	GCTGACAATG	GCATTGGCCT	GACCCCTGGCC	AGTGGTGGAA	CCTTCCCGTA	TGACGACGGC	1440
20	TCCAAGCAAG	AGATAAAGAA	CAGCTTGTTT	GTGGCGGAGA	GTGGCAACGT	GGGGACGGAA	1500
	ATGATGGACA	ATAGGATCTG	GGGCCCTGGC	GGCTTGGACC	ATAGCGGAAG	GACCCCTCCT	1560
	ATAGGCCAGA	ATTTTCCAAT	TAGAGGAAT	CAGTTATATG	ATGGCCCCAT	CAACATCCAA	1620
	AACTGCACTT	TCCGAAAGTT	TGTGGCCCTG	GAGGGCCGGC	ACACCAGCGC	CCTGGCCTTC	1680
	CGCCTGAATA	ATGCCTGGCA	GAGCTGCCCC	CATAACAACG	TGACCGGCAT	TGCCTTTGAG	1740
25	GACGTTCCGA	TTACTTCCAG	AGTGTCTTTC	GGAGAGCCTG	GGCCCTGGTT	CAACCAGCTG	1800
	GACATGGATG	GGGATAAGAC	ATCTGTGTTC	CATGACGTCG	ACGGCTCCGT	GTCCGAGTAC	1860
	CCTGGCTCCT	ACCTCACGAA	GAATGACAAC	TGGCTGGTCC	GGCACCAGCA	CTGCATCAAT	1920
	GTTCCCGACT	GGAGAGGGGC	CATTGTCAGT	GGTGATATG	CACAGATGTA	CATTCAAGCC	1980
	TACAAGACCA	GTAACCTGCG	AATGAAGATC	ATCAAGAATG	ACTTCCCCAG	CCACCTCTT	2040
30	TACCTGGAGG	GGGCGCTCAC	CAGGAGCACC	CATTACCAGC	AATACCAACC	GGTTGTCAAC	2100
	CTGCAGAAGG	GCTACACCAT	CCACTGGGAC	CAGACGGCCC	CCGCCGAAC	CGCCATCTGG	2160
	CTCATCAACT	TCAACAAGGG	CGACTGGATC	CGAGTGGGGC	TCTGCTACCC	GCGAGGCACC	2220
	ACATTCTCCA	TCCTCTCGGA	TGTTCACAAT	CGCCTGCTGA	AGCAAACGTC	CAAGACGGGC	2280
	GTCTTCTGTA	GGACCTTGCA	GATGGACAAA	GTGAGCAGA	GCTACCCTGG	CAGGAGCCAC	2340
35	TACTACTGGG	ACGAGGACTC	AGGGCTGTGT	TTCTGGAAGC	TGAAAGCTCA	GAACGAGAGA	2400
	GAGAAGTTTG	CTTCTGTCTC	CATGAAAGGC	TGTGAGAGGA	TAAAGATTAA	AGCTCTGATT	2460
	CCAAAGAACG	CAGGCGTCAG	TGACTGCACA	GCCACAGCTT	ACCCCAAGTT	CACCGAGAGG	2520
	GCTGTCTGAT	ACGTGCCGAT	GCCCAAGAAG	CTCTTTGGTT	CTCAGCTGAA	AACAAAGGAC	2580
	CATTTCTTGG	AGGTGAAGAT	GGAGAGTTCC	AAGCAGCACT	TCTTCCACCT	CTGGAACGAC	2640
40	TTGCTTACA	TTGAAGTGA	TGGGAAGAGG	TACCCAGTT	CGGAGGATGG	CATCCAGGTG	2700
	GTGGTGATTG	ACGGGAACCA	AGGGCGCGTG	GTGAGCCACA	CGAGCTTCAG	GAACCTCATT	2760
	CTGCAAGGCA	TACCATGGCA	GCTTTTCAAC	TATGTGGCGA	CCATCCCTGA	CAATTCCATA	2820
	GTGCTTATGG	CATCAAAGGG	AAGATACGTC	TCCAGAGGCC	CATGGACCAG	AGTGCTGGAA	2880
	AAGCTTGGGG	CAGACAGGGG	TCTCAAGTTG	AAAGAGCAAA	TGGCATTCGT	TGGCTTCAAA	2940
45	GGCAGCTTCC	GGCCCATCTG	GGTGACACTG	GACACTGAGG	ATCACAAAGC	CAAAATCTTC	3000
	CAAGTTGTGC	CCATCCCTGT	GGTGAAGAAG	AAGAAGTTGT	GAGGACAGCT	GCCGCCCGGT	3060
	GCCACTTCGT	GGTAGACTAT	GACGGTGACT	CTTGGCAGCA	GACCACTGGG	GGATGGCTGG	3120
	GTCCCCCAGC	CCCTGCCAGC	AGCTGCCCTG	GAAGGCCGTG	TTTCAGCCCT	GATGGGCCAA	3180
	GGGAAGGCTA	TCAGAGACCC	TGGTGCTGCC	ACCTGCCCTT	ACTCAAGTGT	CTACCTGGAG	3240
50	CCCCTGGGGC	GGTGCTGGCC	AATGCTGGAA	ACATTCACTT	TCCTGCAGCC	TCTTGGGTGC	3300
	TTCTCTCCTA	TCTGTGCCCT	TTCAGTGGGG	GTTTGGGGAC	CATATCAGGA	GACCTGGGTT	3360
	GTGCTGACAG	CAAAGATCCA	CTTTGGCAGG	AGCCCTGACC	CAGCTAGGAG	GTAGTCTGGA	3420
	GGGCTGGTCA	TTACAGATC	CCCAGGCTCT	TCAGAGACAG	AGTGAGGGTG	GTAAATGTAG	3480
	GAGAAAGAGC	CTTGGCCTTA	AGGAAATCTT	TACTCCTGTA	AGCAAGAGCC	AACCTCACAG	3540
55	GATTAGGAGC	TGGGGTAGAA	CTGGCTATCC	TTGGGGAAGA	GGCAAGCCCT	GCCTCTGGCC	3600
	GTGTCCACCT	TTCAGGAGAC	TTTGAGTGGC	AGGTTTGGAC	TTGGACTAGA	TGACTCTCAA	3660
	AGGCCCTTTT	AGTTCTGAGA	TTCCAGAAAT	CTGCTGCATT	TCACATGGTA	CCTGGAACCC	3720
	AACAGTTCAT	GGATATCCAC	TGATATCCAT	GATGCTGGGT	GCCCCAGCGC	ACACGGGATG	3780
	GAGAGGTGAG	AACTAATGCC	TAGCTTGAGG	GGTCTGCAGT	CCAGTAGGGC	AGGCAGTCAG	3840
60	GTCCATGTGC	ACTGCAATGC	CAGGTGGAGA	AATCACAGAG	AGGTAAATG	GAGGCCAGTG	3900
	CCATTTTACA	GGGGAGGCTC	AGGAAGGCTT	CTTGCTTACA	GGAATGAAGG	CTGGGGGCAT	3960
	TTTGCTGGGG	GGAGATGAGG	CAGCCTCTGG	AATGGCTCAG	GGATTCAGCC	CTCCCTGCCG	4020
	CTGCCTGCTG	AAGCTGGTGA	CTACGGGGTC	GCCCTTTGCT	CACGCTCTCT	TGGCCCACTC	4080
	ATGATGGAGA	AGTGTGGTCA	GAGGGGAGCA	ATGGGCTTTG	CTGCTTATGA	GCACAGAGGA	4140
65	ATTCAGTCCC	CAGGCAGCCC	TGCTCTGAG	TCCAAGAGGG	TGAAGTCCAC	AGAAGTGAGC	4200
	TCCTGCCTTA	GGGCTTCATT	GGCTCTTCAT	CCAGGGAAC	GAGCACAGGG	GGCCTCCAGG	4260
	AGACCCTAGA	TGTGCTCGTA	CTCCCTCGGC	CTGGGATTTT	AGAGCTGGAA	ATATAGAAAA	4320
	TATCTAGCCC	AAAGCCTTCA	TTTAAACAGA	TGGGGAAAGT	GAGCCCCCAA	GATGGGAAAG	4380
	AACCACACAG	CTAAGGGAGG	GCCTGGGGAG	CCCCACCCTA	GCCCTTGCTG	CCACACCACA	4440
70	TTGCCTCAAC	AACCGCCCCC	AGAGTGCCCA	GGCACTCCTG	AGGTAGCTTC	TGGAAATGGG	4500
	GACAAGTCCC	CTCGAAGGAA	AGGAAATGAC	TAGAGTAGAA	TGACAGCTAG	CAGATCTCTT	4560
	CCCTCCTGCT	CCCAGCGCAC	ACAAACCCCG	CCTCCCTTGG	GTGTTGGCGG	TCCCTGTGGC	4620
	CTTCACTTTG	TTCACTACCT	GTCAGCCAG	CCTGGGTGCA	CAGTAGCTGC	AACTCCCAT	4680
	TGGTGCTACC	TGGCTCTCCT	GTCTCTGCAG	CTCTACAGGT	GAGGCCAGC	AGAGGGAGTA	4740
75	GGGCTCGCCA	TGTTTCTGGT	GAGCCAATTT	GGCTGATCTT	GGGTGTCTGA	ACAGCTATTG	4800
	GGTCCACCCC	AGTCCCTTTC	AGCTGCTGCT	TAATGCCCTG	CTCTCTCCCT	GGCCACCTT	4860
	ATAGAGAGCC	CAAAGAGCTC	CTGTAAGAGG	GAGAATCTA	TCTGTGGTTT	ATAATCTTGC	4920
	ACGAGGCACC	AGAGTCTCCC	TGGGTCTTGT	GATGAACTAC	ATTTATCCCC	TTTCTGCCC	4980
	CAACCACAAA	CTCTTTCCTT	CAAAGAGGGC	CTGCCCTGGT	CCCTCCACCC	AACTGCACCC	5040
80	ATGAGACTCG	GTCCAAGAGT	CCATTCCCCA	GGTGGGAGCC	AACTGTCAGG	GAGGTCTTTC	5100
	CCACCAACA	TCTTTCAGCT	GCTGGGAGGT	GACCATAGGG	CTCTGCTTTT	AAAGATATGG	5160
	CTGCTTCAAA	GGCCAGAGTC	ACAGGAAGGA	CTTCTTCCAG	GGAGATTAGT	GGTGATGGAG	5220
	AGGAGAGTTA	AAATGACCTC	ATGTCCTTCT	TGTCCACGGT	TTTGTTGAGT	TTTCACTCTT	5280
	CTAATGCAAG	GGTCTCACAC	TGTGAACCAC	TAGGATGTG	ATCACTTTCA	GGTGGCCAGG	5340
85	AATGTTGAAT	GTCTTTGGCT	CAGTTTCAAT	AAAAAAGATA	TCTATTTGAA	AGTTCTCAGA	5400
	GTTGTACATA	TGTTTCACAG	TACAGGATCT	GTACATAAAA	GTTTCTTTCC	TAAACCATTC	5460
	ACCAAGAGCC	AATATCTAGG	CATTTTCTTG	GTAGCACAAA	TTTTCTTATT	GCTTAGAAAA	5520
	TTGTCTCTCT	TGTTATTCTT	GTTTGTAAGA	CTTAAGTGAG	TTAGGTCTTT	AAGGAAAGCA	5580



ACGCTCCTCT GAAATGCTTG TCTTTTTTCT GTTGCCGAAA TAGCTGGTCC TTTTTCGGGA 5640  
 GTTAGATGTA TAGAGTGTGT GTATGTAAAC ATTTCTTGTA GGCATCACCA TGAACAAAGA 5700  
 TATATTTTCT ATTTATTTAT TATATGTGCA CTTCAAGAAG TCACTGTCAG AGAAATAAAG 5760  
 AATTGTCTTA AATGTCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA

Seq ID NO: 451 Protein sequence  
 Protein Accession #: XP\_051860.2

1 11 21 31 41 51  
 | | | | |  
 MDGVNLSTEV VYKKGQDYRF ACYDRGRACR SYRVRFLCGK PVRPKLTVTI DTNVNSTILN 60  
 LEDNVQSWKP GDTLVIASTD YSMYQAEFQ VLPSCRSCAPN QVKVAGKPMY LHIGEEIDGV 120  
 DMRAEVGLLS RNIIIVMGEME DKCYPYRNHI CNFFDFDTFG GHIKFALGFK AAHLEGTELK 180  
 HMQQLVGGY PIHFHLAGDV DERGGYDPPT YIRDL SIHHT FSRCVTVHGS NGLLIKDVVG 240  
 YNSLGHCFFT EDGPEERNTF DHCLGLLVKS GTLLPSDRDS KMCKMITGDS YPGYIPKPRQ 300  
 DCNAVSTFWM ANPNNNLINC AAAGSEETGF WFIFHHVPTG PSVGMYSFGY SEHIPLGKFY 360  
 NNRAHSNYRA GMIIDNGVKT TEASAKDKRP FLSIISARYS PHQDADPLKP REPAIRHFI 420  
 AYKNQDHGAW LRGGDVWLDL CRFADNGIGL TLASGGTFPY DDGSKQEIKN SLFVGESGNV 480  
 GTEMMDNRIW GPGGLDHSGR TLPFGCNFPI RGIQLYDGPI NIQNCFTFRKF VALEGRHTSA 540  
 LAFRLNNAWQ SCPHNNVTGI AFEDVPITSR VFFGEPGPWF NQLDMGDGKT SVFHDVDGSV 600  
 SEYPGSYLTK NDNWLVRHPD CINVPDWRGA ICSGCYAQMY IQAYKTSNLR MKI IKNDFPS 660  
 HPLYLEGALT RSTHYQQYQP VVTLQKGYTI HWDQTAPAEI AIWLINFNKG DWIRVGLCYP 720  
 RGTTF SILSD VHNRLKQTS KTG VVFTLQ MDKVEQSYPG RSHYYWDEDS GLLFLKLKAQ 780  
 NEREKFAFCS MKGCERIKI ALIPKNAGVS DCTATAYPKF TERAVIDVPM PKKLFGSQLK 840  
 TKDHFLEVKM ESSKQHFFHL WNDFAIYEDV GKYPSSSEDG IQVVVIDGNQ GRVVSHTSFR 900  
 NSILQGIPWQ LFNIVATIPD NSIVLMSKRG RYVSRGPWTR VLEKLGADRG LKLKEQMAFV 960  
 GFKGSFRPIW VTLDTEHKA KIFQVVPFV VKKKKL

Seq ID NO: 452 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 261..2861

1 11 21 31 41 51  
 | | | | |  
 GAGCTAGCGC TCAAGCAGAG CCCAGCGCGG TGCTATCGGA CAGAGCCTGG CGAGCGCAAG 60  
 CGGCGCGGGG AGCCAGCGGG GCTGAGCGCG GCCAGGGTCT GAACCCAGAT TTCCAGACT 120  
 AGCTACCACT CCGCTTGCCC ACGCCCCGGG AGCTCGCGGC GCCTGGCGGT CAGCGACCAG 180  
 ACGTCCGGGG CCGCTGCGCT CCTGGCCCGC GAGGCGTGAC ACTGTCTCGG CTACAGACCC 240  
 AGAGGGAGCA CACTGCCAGG ATGGGAGCTG CTGGGAGGCA GGACTTCCTC TTCAAGGCCA 300  
 TGCTGACCAT CAGCTGGCTC ACTCTGACCT GCTTCCCTGG GGCCACATCC ACAGTGGCTG 360  
 CTGGGTGCCC TGACCAGAGC CCTGAGTTGC AACCTGGAA CCCTGGCCAT GACCAAGACC 420  
 ACCATGTGCA TATCGGCCAG GGCAAGACAC TGCTGCTCAC CTCTTCTGCC ACGGTCTATT 480  
 CCATCCACAT CTCAGAGGGA GGCAAGCTGC TCATTAAAGA CCACGACGAG CCGATTGTTT 540  
 TGCGAACCCT GCACATCTTG ATTGACAACG GAGGAGAGCT GCATGCTGGG AGTGCCCTCT 600  
 GCCCTTTTCCA GGGCAATTTT ACCATCATTT TGTATGGAAG GGCTGATGAA GGTATTTCAGC 660  
 CGGATCCTTA CTATGGTCTG AAGTACATTG GGGTTGGTAA AGGAGGCGCT CTGAGTTGC 720  
 ATGGACAGAA AAAGCTCTCC TGGACATTTT TGAACAAGAC CCTTCACCCA GGTGGCATGG 780  
 CAGAAGGAGG CTATTTTTTT GAAAGGAGCT GGGGCCACCG TGGAGTTATT GTTCATGTCA 840  
 TCGACCCCAA ATCAGGCACA GTCATCCATT CTGACCGGTT TGACACCTAT AGATCCAAGA 900  
 AAGAGAGTGA ACGTCTGGTG CAGTATTTGA ACGCGGTGCC CGATGGCAGG ATCCTTTCTG 960  
 TTGCAAGTGA TGATGAAGGT TCTCGAAATC TGGATGACAT GGCCAGGAAG GCGATGACCA 1020  
 AATTGGGAAG CAAACACTTC CTGCACCTTG GATTTAGACA CCCTTGGAGT TTTCTAACTG 1080  
 TGAAAGGAAA TCCATCATCT TCAGTGAAG ACCATATTGA ATATCATGGA CATCGAGGCT 1140  
 CTGCTGCTGC CCGGCTATTG AAATTGTTCC AGACAGAGCA TGGCGAATAT TTCAATGTTT 1200  
 CTTTGTCCAG TGAGTGGGTT CAAGACGTGG AGTGGACGGA GTGGTTTCGAT CATGATAAAG 1260  
 TATCTCAGAC TAAAGGTGGG GAGAAAATTT CAGACCTCTG GAAAGCTCAC CCAGGAAAAA 1320  
 TATGCAATCG TCCCATTGAT ATACAGGCCA CTACAATGGA TGGAGTTAAC CTCAGCACCG 1380  
 AGGTTGTCTA CAAAAAGGC CAGGATTATA GGTTTGCTTG CTACGACCGG GGCAGAGCCT 1440  
 GCCGGAGCTA CCGTGTACGG TTCCTCTGTG GGAAGCCTGT GAGGCCCAA CTCACAGTCA 1500  
 CCATTGACAC CAATGTGAAC AGCACCATTG TGAACCTGGA GGATAATGTA CAGTCATGGA 1560  
 AACCTGGAGA TACCTGGTTC ATTGCCAGTA CTGATTACTC CATGTACCAG GCAGAAGAGT 1620  
 TCCAGGTGCT TCCCTGCAGA TCCTGCGCCC CCAACAGGT CAAAGTGGA GGGAAACCAA 1680  
 TGTACTGCA CATCGGGGAG GAGATAGACG GCGTGGACAT GCGGGCGGAG GTTGGGCTTC 1740  
 TGAGCCGGAA CATCATAGTG ATGGGGGAGA TGGAGGACAA ATGCTACCCC TACAGAAACC 1800  
 ACATCTGCAA TTTCTTTGAC TTCGATACCT TTGGGGGCCA CATCAAGTTT GCTCTGGGAT 1860  
 TTAAGGCAGC AACTTTGGAG GGCACGGAGC TGAAGCATAT GGGACAGCAG CTGGTGGGTC 1920  
 AGTACCCGAT TCACTTCCAC CTGGCCGGTG ATGTAGACGA AAGGGGAGGT TATGACCCAC 1980  
 CCACATACAT CAGGGACCTC TCCATCCATC ATACATTCTC TCGCTGCGTC ACAGTCCATG 2040  
 GCTCCAATGG CTTGTTGATC AAGGACGTTG TGGGCTATAA CTCTTTGGGC CACTGCTTCT 2100  
 TCACGGAAGA TGGGCCGGAG GAACGCAACA CTTTGTGACC CTGTCTTGGC CTCCTTGTCA 2160  
 AGTCTGGAAC CCTCTCCCC TCAGGACGTC ACAGCAAGAT GTGCAAGATG ATCAGAGAGG 2220  
 ACTCCTACCC AGGGTACATC CCAAGCCCA GGCAAGACTG CAATGCTGTG TCCACCTTCT 2280  
 GGATGGCCAA TCCCAACAAC AACCTCATCA ACTGTGCCGC TGCAGGATCT GAGGAACTG 2340  
 GATTTTGGTT TATTTTTCAC CACGTACCAA CCGGCCCTC CGTGGGAATG TACTCCCCAG 2400  
 GTTATTGAGA GCACATTCCA CTGGGAAAAA TCTATAACA CCGAGCACAT TCCAACCTACC 2460  
 GGGCTGGCAT AATCATAGAC AACCGAGTCA AAACACCGA GGCCTCTGCC AAGGACAAGC 2520  
 GGCCGTTCCT CTCAATCATC TCTGCCAGAT ACAGCCCTCA CCAGGACGCC GACCCGCTGA 2580  
 AGCCCCGGGA GCCGGCCATC ATCAGACACT TCATTGCCTA CAAGAACCAG GACCACGGGG 2640  
 CCTGGCTGCG CCGCGGGGAT GTGTGCTGG ACAGCTGCCA TTTCAGAGGG GAGGCTCAGG 2700  
 AAGGCTTCTT GCTTACAGGA ATGAAGGCTG GGGGCATTTT GCTGGGGGGA GATGAGGCAG 2760  
 CCTCTGGAAT GGCTCAGGGA TTCAGCCCTG CCTGCCCTG CTGCTGAAG CTGGTGACTA 2820  
 CCGGGTCGCC CTTTCTCTAC GTCTCTCTG CCACTCATG ATGGAGAAGT GTGGTCAGAG 2880  
 GGGAGCAATG GGCTTTGCTG CTTATGAGCA CAGAGGAATT CAGTCCCCAG GCAGCCCTGC 2940  
 CTCTGACTCC AAGAGGGTGA AGTCCACAGA AGTGAGCTCC TGCCTTAGGG CCTCATTTGC 3000  
 TCTTCATCCA GGGAACTGAG CACAGGGGGC CTCCAGGAGA CCCTAGATGT GCTCGTACTC 3060  
 CCTCGGCCTG GGATTTCAGA GCTGGAAATA TAGAAAATAT CTAGCCCAA GCCTTCATTT 3120



5  
10  
15  
20  
25

TAACAGATGG	GGAAAGTGAG	CCCCCAAGAT	GGGAAAGAAC	CACACAGCTA	AGGGAGGGCC	3180
TGGGGAGCCC	CACCCTAGCC	CTTGCTGCCA	CACCACATTG	CCTCAACAAC	CGGCCCCAGA	3240
GTGCCCAGGC	ACTCCTGAGG	TAGCTTCTGG	AAATGGGGAC	AAGTCCCCTC	GAAGGAAAGG	3300
AAATGACTAG	AGTAGAATGA	CAGCTAGCAG	ATCTCTTCCC	TCCTGCTCCC	AGCGCACACA	3360
AACCCGCCCT	CCCCTTGGTG	TTGGCGGTCC	CTGTGGCCTT	CACCTTGTTC	ACTACCTGTC	3420
AGCCCAGCCT	GGGTGCACAG	TAGCTGCAAC	TCCCCATTGG	TGCTACCTGG	CTCTCCTGTC	3480
TCTGCAGCTC	TACAGGTGAG	GCCCAGCAGA	GGGAGTAGGG	CTCGCCATGT	TTCTGGTGAG	3540
CCAATTTGGC	TGATCTTGGG	TGTCTGAACA	GCTATTGGGT	CCACCCCAGT	CCCTTTTCAGC	3600
TGCTGCTTAA	TGCCCTGCTC	TCTCCCTGGC	CCACCTTATA	GAGAGCCCAA	AGAGCTCCTG	3660
TAAGAGGGAG	AACTCTATCT	GTGGTTTATA	ATCTTGCACG	AGGCACCAGA	GTCTCCCTGG	3720
GTCTTGTGAT	GAACCTACATT	TATCCCCTTT	CCTGCCCCAA	CCACAAACTC	TTTCCTTCAA	3780
AGAGGGCCTG	CCTGGCTCCC	TCCACCCAAC	TGCACCCATG	AGACTCGGTC	CAAGAGTCCA	3840
TTCCCAGGT	GGGAGCCAAC	TGTCAGGGAG	GTCTTTCCCA	CCAAACATCT	TTTCTGCTGCT	3900
GGGAGGTGAC	CATAGGGCTC	TGCTTTTAAA	GATATGGCTG	CTTCAAAGGC	CAGAGTCACA	3960
GGAAGGACTT	CTTCCAGGGA	GATTAGTGGT	GATGGAGAGG	AGAGTTAAAA	TGACCTCATG	4020
TCCTTCTTGT	CCACGGTTTT	GTGAGTTTTT	CACTCTTCTA	ATGCAAGGGT	CTCACACTGT	4080
GAACCACTTA	GGATGTGATC	ACTTTCAGGT	GGCCAGGAAT	GTTGAATGTC	TTTGGCTCAG	4140
TTCATTTAAA	AAAGATATCT	ATTTGAAAGT	TCTCAGAGTT	GTACATATGT	TTTACAGTAC	4200
AGGATCTGTA	CATAAAAGTT	TCTTTCTCTA	ACCATTACAC	AAGAGCCAAT	ATCTAGGCAT	4260
TTTCTTGGTA	GCACAAATTT	TCTTTATGCT	TAGAAAATTG	TCCTCCTTGT	TATTTCTGTT	4320
TGTAAGACTT	AAGTGAGTTA	GGTCTTTAAG	GAAAGCAACG	CTCCTCTGAA	ATGCTTGTCT	4380
TTTTTCTGTT	GCCGAAATAG	CTGGTCCTTT	TTCCGGAGTT	AGATGTATAG	AGTGTGTTGTA	4440
TGTAAACATT	TCTTGTAGGC	ATCACCATGA	ACAAAGATAT	ATTTTCTATT	TATTTATTAT	4500
ATGTGCACTT	CAAGAAGTCA	CTGTCAGAGA	AATAAAGAAT	TGTCTTAAAT	GTCTATGATTG	4560
GAGATGTCTT	TTGCATTGCT	TGGAAGGGGT	GTACCTAGAG	CCAAGGAAAT	TGGCTCTGGT	4620
TTTGGAAAAA	TTTGCTGTTA	TTATAGTAAA	CATACAAAGG	ATGTCAAAAA	AAAAAAAAAA	4680
AAAAAAAAAA	AAAAAAAAAA	AA				

Seq ID NO: 453 Protein sequence  
Protein Accession #: Eos sequence

35  
40  
45

1	11	21	31	41	51	
MGAAGRQDFL	FKAMLTISWL	TLTCFPGATS	TVAAGCPDQS	PELQPNWPGH	DQDHHVHIGQ	60
GKTLTLLTSSA	TVYSIHISEG	GKLVIKDHDE	PIVLRTRHIL	IDNGGELHAG	SALCPFQGNF	120
TIILYGRADE	GIQDPYLYGL	KYIGVGKGGG	LELHGQKKLS	WTFNLKTLHP	GGMAEGGYFF	180
ERSWGHGRVI	VHVIDPKSGT	VIHSDRFDTY	RSKKESERLV	QYLNAPVDGR	ILSVAVNDEG	240
SRNLDDMARK	AMTKLGSKHF	LHLGFRHPWS	FLTVKGNPSS	SVEDHIEYHG	HRGSAAARVF	300
KLFQTEHGEY	FNVSLSSEWV	QDVEWTEWFD	HDKVSQTKGG	EKISDLWKAH	PGKICNRPID	360
IQATTMDGVN	LSTEVVYKKG	QDYRFACYDR	GRACRSYRVR	FLCGKPVPRK	LTVTIDTNVN	420
STILNLEDNV	QSWKPGDTLV	IASTDYSMYQ	AEEFQVLPCR	SCAPNQVKVA	GKPMYLIHIG	480
EIDGVDMRAE	VGLLSRNIIIV	MGEMEDKCYP	YRNHICNFFD	FDTFGGHIKF	ALGFKAAHLE	540
GTELKHMGGQ	LVGQYPIHFH	LAGDVDERGG	YDPPTYIRDL	SIHHTFSRCV	TVHGSNGLLI	600
KDVGYNLSLG	HCFFTEDGPE	ERNTFDHCLG	LLVKSGTLLP	SDRDSKMCKM	ITEDSYPGYI	660
PKPRQDCNAV	NLINCMAAGS	EETGFWFIFH	HVPTGPSVGM	YSPGYSEHIP		720
LGKFNNNRAH	SNYRAGMIID	NGVKTTEASA	KDKRPFLSII	SARYSPHQDA	DPLKPREPAI	780
IRHFIAYKNQ	DHGAWLRGGD	VWLDSCHFRR	EAQEGFLLTG	MKAGGILLGG	DEAASGMAQG	840
FSPPCRCLLK	LVTGTGSPFAH	VSLAHS				

Seq ID NO: 454 DNA sequence  
Nucleic Acid Accession #: NM\_013282.2  
Coding sequence: 85..2466

55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
CGACTCCTTA	GAGCATGGCA	TGGCTCAGAG	GTGCTGGTAA	AACTGATGGG	GGTTTTTGCT	60
GTCCCTCCCC	TCAGCGCCGA	CACCATGTGG	ATCCAGGTTT	GGACCATGGA	CGGGAGGCAG	120
ACCCACACGG	TGGACTCGCT	GTCCAGGCTG	ACCAAGGTGG	AGGAGCTGAG	GCGGAAGATC	180
CAGGAGCTGT	TCCACGTGGA	GCCAGGCTTG	CAGAGGCTGT	TCTACAGGGG	CAAACAGATG	240
GAGGACGGCC	ATACCCTCTT	CGACTACGAG	GTCCGCTGTA	ATGACACCAT	CCAGCTCCTG	300
GTCCGCCAGA	GCCTCGTGCT	CCCCACAGC	ACCAAGGAGC	GGGACTCCGA	GCTCTCCGAC	360
ACCGACTCCG	GCTGCTGCC	GGGCCAGAGT	GAGTCAGACA	AGTCTCCAC	CCACGGCGAG	420
GCGGCCGCCG	AGACTGACAG	CAGGCCAGCC	GATGAGGACA	TGTGGGATGA	GACGGAATTG	480
GGGCTGTACA	AGGTCAATGA	GTACGTGCGT	GCTCGGGACA	CGAACATGGG	GGCGTGGTTT	540
GAGGCGCAGG	TGGTCAGGGT	GACGCGGAAG	GCCCCCTCCC	GGGACGAGCC	CTGCAGCTCC	600
ACGTCCAGGC	CGGCGCTGGA	GGAGGACGTC	ATTTACCACG	TGAAATACGA	CGACTACCCG	660
GAGAACGGCG	TGGTCCAGAT	GAATCCAGG	GACGTCCGAG	CGCGCGCCCG	CACCATCATC	720
AAGTGGCAGG	ACCTGGAGGT	GGGCCAGGTG	GTCATGCTCA	ACTACAACCC	CGACAACCCC	780
AAGGAGCGGG	GCTTCTGGTA	CGACGCGGAG	ATCTCCAGGA	AGCGCGAGAC	CAGGACGGCG	840
CGGGAACCTC	ACGCCAACGT	GCTGCTGGGG	GGTGATTCTC	TGAACGACTG	TCGGATCATC	900
TTCTGTGACG	AAGTCTTCAA	GATTGAGCGG	CCGGGTGAAG	GGAGCCCCAT	GGTTGACAAC	960
CCCATGAGAC	GGAAGAGCGG	GCCGTCTGTC	AAGCACTGCA	AGGACGACGT	GAACAGACTC	1020
TGCCGGGTCT	GCGCCTGCCA	CCTGTGCGGG	GGCCGGCAGG	ACCCCGACAA	GCAGCTCATG	1080
TGCGATGAGT	GCGACATGGC	CTTCCACATC	TACTGCCTGG	ACCCGCCCTT	CAGCAGTGTT	1140
CCCAGCGAGG	ACGAGTGGTA	CTGCCCAATG	TGCCGGAATG	ATGCCAGCGA	GGTGGTACTG	1200
GCGGGGAGAG	GGCTGAGAGA	GAGCAAGAAG	AGGCCGAAGA	TGGCCTCGGC	CACATCGTCC	1260
TCACAGCGGG	ACTGGGGCAA	GGGCATGGCC	TGTGTGGGCC	GCACCAAGGA	ATGTACCATC	1320
GTCCCGTCCA	ACCACTACGG	ACCCATCCCG	GGGATCCCCG	TGGGCACCAT	GTGGCGGTTT	1380
CGAGTCCAGG	TCAGCGAGTC	GGGTGTCCAT	CGGCCCCACG	TGGCTGGCAT	ACACGGCCGG	1440
AGCAACGACG	GAGCGTACTC	CCTAGTCTTG	GCGGGGGGCT	ATGAGGATGA	CGTGGACCAT	1500
GGGAATTTTT	TCACATACAC	GGGTAGTGGT	GGTCGAGATC	TTTCCGGCAA	CAAGAGGACC	1560
GCGGAACAGT	CTTGTGATCA	GAAACTCACC	AACACAACA	GGGCGCTGGC	TCTCAACTGC	1620
TTTGCTCCCA	TCAATGACCA	AGAAGGGGCC	GAGGCCAAGG	ACTGGCGGTC	GGGGAAGCCG	1680
GTCAGGGTGG	TGCGCAATGT	CAAGGGTGGC	AAGAATAGCA	AGTACGCCCC	CGCTGAGGGC	1740
AACCGCTACG	ATGGCATCTA	CAAGGTGTGT	AAATACTGGC	CCGAGAAGGG	GAAGTCCGGG	1800
TTTCTCGTGT	GGCGCTACCT	TCTGCGGAGG	GACGATGATG	AGCCTGGCCC	TTGGACGAAG	1860
GAGGGGAAGG	ACCGGATCAA	GAAGCTGGGG	CTGACCATGC	AGTATCCAGA	AGGCTACCTG	1920



5  
10  
15  
20  
25  
30  
35

GAAGCCCTGG CCAACCGAGA GCGAGAGAAG GAGAACAGCA AGAGGGAGGA GGAGGAGCAG 1980  
CAGGAGGGGG GCTTCGCGTC CCCCAGGACG GGCAAGGGCA AGTGAAGCG GAAGTCGGCA 2040  
GGAGGTGGCC CGAGCAGGGC CGGGTCCCCG CGCCGGACAT CCAAGAAAAC CAAGGTGGAG 2100  
CCCTACAGTC TCACGGCCCA GCAGAGCAGC CTCATCAGAG AGGACAAGAG CAACGCCAAG 2160  
CTGTGGAATG AGGTCCTGGC GTCACCTAAG GACCGGCCGG CGAGCGGCAG CCCGTTCAG 2220  
TTGTTCTCTGA GTAAAGTGA GGAGACGTTT CAGTGTATCT GCTGTCAGGA GCTGGTGTTC 2280  
CGGCCCATCA CGACCGTGTG CCAGCACAAC GTGTGCAAGG ACTGCCTGGA CAGATCCTTT 2340  
CGGGCACAGG TGTTCACTG CCCTGCCTGC CGCTACGACC TGGGCCGCAG CTATGCCATG 2400  
CAGGTGAACC AGCCTCTGCA GACCGTCTCT AACCAGCTCT TCCCCGGCTA CGGCAATGGC 2460  
CGGTGATCTC CAAGCACTTC TCGACAGGCG TTTTGCTGAA AACGTGTCGG AGGGCTCGTT 2520  
CATCGGCACT GATTTTGTTC TTAGTGGGCT TAACCTAAAC AGGTAGTGTT TCCTCCGTTT 2580  
CCTAAAAGG TTTGTCTTCC TTTTTTTTTT TTTTATTTT TCAAACTAT ACATTTTCAG 2640  
GAATTTATGT ATTCTGGCTA AAAGTTGGAC TTCTCAGTAT TGTGTTAGT TCTTTGAAAA 2700  
CATAAAAGCC TGCAATTTCT CGACAAAACA ACACAAGATT TTTTAAAGAT GGAATCAGAA 2760  
ACTACGTGGT GTGGAGGCTG TTGATGTTTC TGGTGTCAAG TTCTCAGAAG TTGCTGCCAC 2820  
CAACTCTTTA AGAAGGCGAC AGGATCAGTC CTTCTCTAGG GTTCTGGCCC CCAAGGTCAG 2880  
AGCAAGCATC TTCCTGACAG CATTTTGTCA TCTAAAGTCC AGTGACATGG TTCCCCGTGG 2940  
TGGCCCGTGG CAGCCCGTGG CATGGCGTGG CTCAGCTGTC GTTGAAGTT GTTGCAAGGA 3000  
AAAGAGGAAA CATCTCGGGC CTAGTTCAAA CCTTGCCTC AAAGCCATCC CCCACCAGAC 3060  
TGCTTAGCGT CTGAGATCCG CGTGAAAAGC CCTTGCCCA CGAGAGCAGG GAGTTGGGGC 3120  
CACGCAGAAA TGGCCTCAAG GGGACTCTGC TCCACGTGGG GCCAGGCGTG TGACTGACGC 3180  
TGTCCGACGA AGGCGGCCAC GGACGGACGC CAGCACACGA AGTCAGTGC AAGTGCCTTT 3240  
GATTGCTTCC TTCTTCTTAA AGACGACAGT CTTTGTGTGT AGCACTGAAT TATTGAAAAT 3300  
GTCAACCAGA TTCTAGAAAC TGCGGTCATC CAGTCTCTCC TGACACCGGA TGGGTGCTTG 3360  
GGAACCGTTT GAGCCTTATA GATCATTTAC ATTCAATTTT TTTAACTCAG CAAGTGAGAA 3420  
CTTACAAGAG GGTTTTTTT TAATTTTTT TTCTCTTAAT GAACACATTT TCTAAATGAA 3480  
TTTTTTTTTGT AGTTACTGTA TATGTACCAA GAAAGATATA ACGTTAGGGT TTGGTTGTTT 3540  
TTGTTTTTGT ATTTTTTTTC TTTTGAAAGG GTTTGTAAAT TTTTCTAATT TTACCAAAGT 3600  
TTGCAGCCTA TACCTCAATA AAACAGGGAT ATTTTAAATC ACATACCTGC AGACAAACTG 3660  
GAGCAATGTT ATTTTAAAG GGTTTTTTTC ACCTCCTTAT TCTTAGATTA TTAATGTATT 3720  
AGGGAAGAAT GAGACAATT TGTGTAGGCT TTTTCTAAAG TCCAGTACTT TGTCCAGATT 3780  
TTAGATTCTC AGAATAAATG TTTTTCACAG ATTGAAAAA AAAAAAA

Seq ID NO: 455 Protein sequence  
Protein Accession #: NP\_037414.2

40  
45  
50

1 11 21 31 41 51  
MWIQVRTMDG RQHTVDSL RLTKVEELRR KIQLFHVPE GLQRLFYRGK QMEDGHTLFD 60  
YEVRLNDTIQ LLVRQSLVLP HSTKERDSEL SDTDSGCCLG QSESDKSSTH GEAAAETDSR 120  
PADEDMWDET ELGLYKVNEY VDARDTNMGA WFEAQVVRVT RKAPSRDEPC SSTSRPALEE 180  
DVIYHVKYDD YPENGVVQMN SRDVRARART IIKWQDLEVG QVVMNLNYPD NPKERGFWD 240  
AEISRKRETR TARELYANVV LGDDSLNDCR IIFVDEVFKI ERPGEPSMV DNPMMRKS GP 300  
SCKHCKDDVN LCRVCACHL CGGRQDPDKQ LMCDECDMAF HIYCLDPPLS SVPSEDEWYC 360  
PECRNDASEV VLAGERLRES KKKAKMASAT SSSQRDWGKG MACVGRTEK TIVPSNHYGP 420  
IPGIPVGTMW RFRVQVSESG VHRPHVAGIH GRSNDGAYSL VLAGGYEDDV DHGNFFTYTG 480  
SGGRDLSGNK RTAEQSCDQK LTNTNRALAL NCFAPINDQE GAEAKDWRSG KPVVRVNRVK 540  
GGKNSKYAPA EGNRYDGIYK VVKYWPEKKG SGFLVWRYLL RRDDDEPGPW TKEGKDRIKK 600  
LGLTMQYPEG YLEALANRER EKENSKREEE EQQEGGFASP RTGKGKWKRK SAGGGPSRAG 660  
SPRRTSKKTK VEPYSLTAQQ SSLIREDKSN AKLWNEVLAS LKDRPASGSP FQLFLSKVEE 720  
TFQCICCCQEL VFRPITTVCC HNVCKDCLDR SFRAQVFSCP ACRYDLGRSY AMQVNQPLQT

Seq ID NO: 456 DNA sequence  
Nucleic Acid Accession #: NM\_001200.1  
Coding sequence: 325..1514

60  
65  
70  
75  
80  
85

1 11 21 31 41 51  
GGGGAATTCT TGAACCTGCA GGGAGAATAA CTTGCGCACC CCACTTTGCG CCGGTGCCTT 60  
TGCCCCAGCG GAGCCTGCTT CGCCATCTCC GAGCCCCACC GCCCTCCAC TCCTCGGCCT 120  
TGCCCCAGAC TGAGACGCTG TTCCAGCGT GAAAAGAGAG ACTGCGCGGC CGGCACCCGG 180  
GAGAAGGAGG AGGCAAAGAA AAGGAACGGA CATTGCGTCC TTGCGCCAGG TCCTTTGACC 240  
AGAGTTTTTC CATGTGACG CTCTTTCAAT GGAGCTGTCC CCGGTGCTT CTAGACGGA 300  
CTGCGGTCTC CTAAAGTCCG ACCATGGTGG CCGGGACCCG CTGTCTTCTA GCGTTGCTGC 360  
TTCCCCAGGT CCTCCTGGGC GGCGCGGCTG GCCTCGTTCC GGAGCTGGGC CGCAGGAAGT 420  
TCGCGGCGGC GTCGTCGGGC CGCCCTCAT CCCAGCCCTC TGACGAGGTC CTGAGCGAGT 480  
TCGAGTTGCG GCTGCTCAGC ATGTTCCGCC TGAAACAGAG ACCCACCACC AGCAGGGACG 540  
CCGTGGTGCC CCCCTACATG CTAGACCTGT ATCGCAGGCA CTCAGGTCAG CCGGGCTCAC 600  
CCGCCCCAGA CCACCGGTTG GAGAGGGCAG CCAGCCGAGC CAACACTGTG CGCAGCTTCC 660  
ACCTGAAGA ATCTTTGGAA GAACAGTACC AAACGAGTGG GAAAAACAACC CGGAGATTCT 720  
TCTTTAATTT AAGTTCTATC CCCACGGAGG AGTTTATCAC CTCAGCAGAG CTTCAGGTTT 780  
TCCGAGAACA GATGCAAGAT GCTTTAGGAA ACAATAGCAG TTTCCATCAC CGAATTAATA 840  
TTTATGAAAT CATAAAACCT GCAACAGCCA ACTCGAAAT CCCCCTGACC AGACTTTTGG 900  
ACACCAGGTT GGTGAATCAG AATGCAAGCA GGTGGGAAAG TTTTGATGTC ACCCCCGCTG 960  
TGATGCGGTG GACTGCACG GGACAGGCA ACCATTGATT CGTGGTGGAA GTGGCCCACT 1020  
TGGAGGAGAA ACAAGGTGTC TCCAAGAGAC ATGTAGGAT AAGCAGGTCT TTGCACCAAG 1080  
ATGAACACAG CTGGTCACAG ATAAGGCCAT TGCTAGTAAC TTTTGGCCAT GATGGAAAAG 1140  
GGCATCTCT CCACAAAAGA GAAAAACGTC AAGCCAAACA CAAACAGCGG AAACGCCTTA 1200  
AGTCCAGCTG TAAGAGACAC CCTTTGTACG TGGACTTCAG TGACGTGGGG TGGAATGACT 1260  
GGATTGTGGC TCCCCCGGGG TATCACGCCT TTTACTGCCA CGGAGAAATGC CCTTTTCTCT 1320  
TGGCTGATCA TCTGAATCC ACTAATCAT CCATTGTTCA GACGTTGGTC AACTCTGTTA 1380  
ACTCTAAGAT TCCTAAGGCA TGCTGTGTCC CGACGAACT CAGTGCTATC TCGATGCTGT 1440  
ACCTTGACGA GAATGAAAAG GTTGTATTAA AGAACTATCA GGACATGGTT GTGGAGGGTT 1500  
GTGGGTGTCT TAGTACAGC AAAATTAAAT ACATAAATAT ATATATA

Seq ID NO: 457 Protein sequence  
Protein Accession #: NP\_001191.1



	1	11	21	31	41	51	
5	MVAGTRCLLA	LLLPOVLLGG	AAGLVPELGR	RKFAAASSGR	PSSQPSDEVL	SEFELRLLSM	60
	FGLKQRPTPS	RDVVPPPYML	DLYRRHSGQP	GSPAPDHRLE	RAASRANTVR	SFHHEESLEE	120
	LPETSGKTTR	RFFFNLSIP	TEEFITSAEL	QVFREQMQDA	LGNNSSFHHR	INIYEI IKPA	180
	TANSKFPVTR	LLDT					

Seq ID NO: 458 DNA sequence  
 Nucleic Acid Accession #: NM\_001999.2  
 Coding sequence: 1..8736

	1	11	21	31	41	51	
15	ATGGGGAGAA	GACGGAGGCT	GTGTCTCCAG	CTCTACTTCC	TGTGGCTGGG	CTGTGTGGTG	60
	CTCTGGGCGC	AGGGCACGGC	CGGCCAGCCT	CAGCCTCCTC	CGCCCAAGCC	GCCCCGGCCC	120
	CAGCCGCCGC	CGCAACAGGT	TCGGTCCGCT	ACAGCAGGCT	CTGAAGCGCG	GTTTCTAGCG	180
	CCCGAGTATC	GCGAGGAGGG	TGCCGCAGTG	GCCAGCCGCG	TCCGCCGGCG	AGGACAGCAG	240
	GACGTGCTCC	GAGGGCCCAA	CGTGTGCGGC	ACTCCTACTG	CTGCCCTGGA		300
20	TGGAAGACGC	TCCCTGGAGG	AAACCAAGTG	ATTGTCCCGA	TTTGTAGAAA	TAGTTGTGGA	360
	GATGGATTTT	GTTCCCGTCC	TAACATGTGT	ACTTGTTCCT	GTGGGCAAAAT	ATCATCAACC	420
	TGTGGATCAA	AATCAATTCA	GCAGTGCAGT	GTGAGATGCA	TGAATGGTGG	GACCTGTGCA	480
	GATGACCACT	GCCAGTGCCA	GAAAGGATAT	ATTGGAACCT	ATTGTGGACA	ACCTGTCTGT	540
	GAAAATGGAT	GTCAGAAATG	TGGACGTTGC	ATCGCCCCAA	CGTGTGCTTG	TGTTTATGGG	600
25	TTCCTGGTTC	CACAGTGTGA	AAGAGATTAC	AGGACAGGCC	CGTGTTCAC	TCAGGTCAAC	660
	AACAGATGT	GCCAAGGGCA	GCTGACAGGC	ATTGTCTGCA	CGAAGACTCT	GTGCTGTGCC	720
	ACCACTGGAC	GGGCGTGGGG	CCATCCCTGT	GAGATGTGTC	CAGCCCAGCC	TCAGCCCTGC	780
	CGACGGGGTT	TCATCCCCAA	CATCCGCACT	GGAGCTTGCC	AAGATGTTGA	TGAATGCCAG	840
	GCTATCCCAG	GGATATGCCA	AGGAGGAAAC	TGTATCAATA	CAGTGGGCTC	TTTTGAATGC	900
30	AGATGCCCTG	CTGGTCACAA	ACAGAGTGAA	ACTACTCAGA	AATGTGAAGA	CATTGATGAG	960
	TGCAGCATCA	TTCCTGGGAT	ATGTGAAACT	GGTGAATGTT	CCAACACCGT	GGGAAGCTAT	1020
	TTTTGTGTTT	GTCCACGTGG	ATATGTAACC	TCAACAGATG	GCTCTCGATG	CATCGATCAG	1080
	AGAACAGGCA	TGTGTTTCTC	GGGCCTGGTG	AATGGCCGCT	GTGCACAAGA	GCTCCCGGGG	1140
	AGAATGACGA	AAATGCAGTG	CTGCTGTGAG	CCTGGCCGCT	GCTGGGGCAT	CGGAACCATT	1200
35	CCTGAAGCCT	GTCCTGTCAG	AGGTTCTGAG	GAATATCGCA	GACTTTGCAT	GGATGGACTT	1260
	CCAATGGGAG	GAATTCCAGG	GAGTGCTGGT	TCCAGACCTG	GAGGCACTGG	GGGAAATGGC	1320
	TTTGCCCCAA	GTGGCAATGG	CAATGGCTAT	GGCCCAGGAG	GGACAGGCTT	CATCCCCATC	1380
	CCTGGAGGCA	ATGGCTTTTC	TCCTGGCGTT	TGGGGAGCCG	GTGTGGGGGC	CGGGGGACAG	1440
	GGACCTATCA	TCACCTGGACT	AACAATTCTG	AACCAGACAA	TAGATATCTG	TAAGCATCAT	1500
40	GCTAACCTTT	GTTTAAATGG	ACGCTGTATA	CCAACTGTCT	CAAGCTACCG	ATGTGAATGC	1560
	AACATGGGTT	ATAAGCAGGA	TGCAAAATGA	GATTGTATAG	ATGTTGATGA	ATGCACATCA	1620
	AATCCCTGCA	CTAATGGAGA	TTGTGTTAAC	ACACCTGGTT	CCTATTATTG	TAAATGTCAT	1680
	GCTGGATTCC	AGAGGACTCC	TACCAAGCAA	GCATGCATTG	ATATTGATGA	GTGCATCCAG	1740
	AATGGGGTTC	TTTGTAAAAA	CGGTCGATGC	GTGAACTCAG	ATGGAAAGTTT	CCAGTGCATT	1800
45	TGCAATGCCG	GCTTTGAATT	AACTATCAGT	GGAAAAAACT	GTGTTGATCA	TGATGAATGT	1860
	ACAACCTACCA	ACATGTGTTT	GAATGGAATG	TGCATCAATG	AAGATGGCAG	CTTCAAGTGC	1920
	ATCTGCAAAC	CAGGATTTGT	CTTGGCTCCA	AATGGGCGTT	ACTGTACTGA	TGTTGATGAA	1980
	TGCCAGACCC	CAGGAATCTG	CATGAATGGG	CACTGCATCA	ACAGTGAAGG	GTCCTTCCGC	2040
	TGTGACTGTC	CCCCAGGCCT	GGCTGTGGGC	ATGGATGGAC	GTGTGTGTGT	TGATACTCAC	2100
50	ATGCGCAGTA	CCTGCTATGG	AGGAATCAAG	AAAGAGTGT	GTGTGCGTCC	TTTCCCCGGT	2160
	GCAGTGACCA	AGTCCGAATG	CTGCTGTGCC	AATCCAGACT	ATGTTTGTGG	AGAACCCTGC	2220
	CAGCCATGCC	CTGCAAAAAA	TTTCACTGAA	TTCCACGGCC	TTTGTAGTAG	TGGAGTAGGT	2280
	ATCACTGTGG	ATGGAAGAGA	TATCAATGAA	TGTGCTTTGG	ATCCTGATAT	ATGTGCCAAT	2340
	GGGATTTGTG	AAAACCTACG	TGGTAGTTAC	CGTTGTAATT	GCAACAGTGG	CTATGAACCA	2400
55	GATGCCTCTG	GAAGAAACTG	TATTGACATT	GATGAATGTT	TAGTAAACAG	ACTGCTTTGT	2460
	GATAACGGAT	TGTGCCGAAA	CACGCCAGGA	AGTTACAGCT	GTACGTGCCC	ACCAGGGTAT	2520
	GTGTTTCAGGA	CTGAGACAGA	GACCTGTGAA	GATATAATG	AATGTGAAAG	CAACCCATGT	2580
	GTCAATGGGG	CCTGCAGAAA	CAACCTTGGA	TCTTTCAATT	GTGAATGTTT	GCCCCGCAGC	2640
	AAACTCAGCT	CCACAGGATT	GATCTGTATT	GACAGCCTGA	AGGGGACCTG	TTGGCTCAAC	2700
60	ATCCAGGACA	GCCGCTGTGA	GGTGAATATT	AATGGAGCCA	CTCTGAAATC	TGAATGCTGT	2760
	GCCACCCTCG	GAGCCGCTCG	GGGGAGCCCC	TGTGAGCGGT	GTGAACCTAG	TACAGCTTGC	2820
	CCAAGAGGGC	TTGCCAGGAT	TAAAGGTGTT	ACGTGTGAAG	ATGTTAATGA	GTGTGAGGTG	2880
	TTCCCTGGCG	TTTGTCCAAA	TGGACGCTGT	GTCAACAGTA	AGGGATCTTT	TCATTGCGAG	2940
	TGCTGCTGAG	GCCTTACGTT	GGATGGGACT	GGCCGTGTAT	GTTTGGATAT	TCGCATGGAG	3000
65	CAGTGTACT	TGAAGTGGGA	TGAAGATGAA	TGCATCCACC	CCGTCTCTGG	AAAGTTCCGC	3060
	ATGGATGCCT	GCTGCTGTGC	TGTCGGGGCG	GCTTGGGGCA	CCGAGTGTGA	GGAGTGCCCC	3120
	AAACCTGGCA	CCAAGGAATA	CGAGACACTG	TGCCCCCGCG	GGGCTGGCTT	TGCTAACCGA	3180
	GGGGATGTTT	TTACTGGGCG	GCCATTTTAC	AAAGACATCA	ATGAATGCAA	AGCATTTCCT	3240
	GGGATGTGCA	CTTATGGGAA	GTGCAGAAAT	ACAATCGGAA	GCTTCAAATG	CCGTTGCAAT	3300
70	AGTGGCTTTG	CTCTAGACAT	GGAGGAAAGA	AACCTGCACG	ACATCGACGA	GTGCAGGATT	3360
	TCTCCTGACC	TCTGTGGCAG	TGGAATCTGC	GTCAATACAC	CGGGCAGCTT	TGAGTGCGAG	3420
	TGCTTCGAAG	GCTATGAAAG	TGGCTTCATG	ATGATGAAGA	ACTGCATGGA	CATTGACGGA	3480
	TGTGAACGTA	ACCCTCTCCT	TTGTAGGGGT	GGCACCTGTG	TGAACACTGA	GGGCAGCTTT	3540
	CAGTGTGACT	GCCCCACTGG	ACACGAGCTG	TCACCATCCC	GTGAGGACTG	TGTGGATATT	3600
75	AATGAATGCT	CCCTGAGTGA	CAATCTCTGC	AGAAATGGAA	AATGTGTGAA	CATGATTGGA	3660
	ACCTATCAGT	GCTCTTGCAA	TCCTGGATAT	CAGGCTACGC	CAGACCGCCA	GGGCTGTACA	3720
	GATATTGATG	AATGTATGAT	AATGAACGGA	GGCTGTGACA	CCCAGTGCAC	AAATTGAGAG	3780
	GGAAGCTACG	AATGCAGCTG	CAGTGAGGGT	TATGCCCTGA	TGCCAGATGG	GAGATCGTGT	3840
	GCAGACATTG	ATGAATGTGA	AAACAATCCT	GATATCTGTG	ATGGCGGCCA	GTGTACCAAC	3900
80	ATTCTGAGAG	AGTATCGCTG	CCTCTGCTAT	GATGGCTTCA	TGGCTTCCAT	GGACATGAAA	3960
	ACATGCATTG	ATGTCAATGA	ATGTGACCTA	AATTCAAATA	TCTGCATGTT	TGGGGAATGT	4020
	GAGAACACAA	AGGGATCCTT	CATTTGCCAC	TGTACAGCTG	GTTACTCAGT	GAAGAAGGGG	4080
	ACCACAGGAT	GTACAGATGT	GATGAGTGTG	GAAATTGGTG	CTCATAACTG	CGACATGCAT	4140
	GCCTCATGTC	TGAATATCCC	AGGAAGCTTC	AAGTGTAGCT	GCAGAGAAGG	CTGGATTGGA	4200
85	AACGGCATCA	AGTGTATTGA	TCTGGACGAA	TGTTCTAATG	GAACCCACCA	GTGTAGCATC	4260
	AATGCTCAGT	GTGTAAATAC	CCCGGGCTCA	TACCGCTGTG	CCTGCTCCGA	AGGTTTCACT	4320
	GGTGATGGCT	TTACCTGCTC	AGATGTTGAT	GAGTGTGCAG	AAAACATAAA	CCTCTGTGAG	4380



	AACGGACAGT	GCCTTAATGT	CCCGGGTGCA	TATCGCTGCG	AGTGTGAGAT	GGGCTTCACT	4440
	CCAGCCTCAG	ACAGCAGATC	CTGCCAAGAT	ATTGATGAAT	GCTCCTTCCA	AAACATTTGT	4500
	GTCTCTGGAA	CATGTAATAA	CTGCGCTGGA	ATGTTTCATT	GCATCTGCGA	TGATGGTTAT	4560
5	GAATTGGACA	GAACAGGAGG	GAACGTGTACA	GATATTGATG	AGTGTGCAGA	TCCTATAAAC	4620
	TGTGTCAATG	GCCTATGTGT	CAACACGCCT	GGTCGCTATG	AGTGTAACTG	CCCACCCGAT	4680
	TTTCAGTTGA	ACCCAACCTG	TGTGGGTTGT	GTTGACAACC	GTGTGGGCAA	CTGCTACCTG	4740
	AAGTTTGGAC	CTCGAGGAGA	TGGGAGTCTG	TCTTGCAACA	CCGAGATCGG	GGTGGGCGTC	4800
	AGTCGCTCTT	CATGCTGCTG	CTCTCTGGGA	AAGGCCCTGG	GAAACCCCTG	TGAGACATGC	4860
10	CCCCCTGTCA	ATAGCACTGA	ATATTACACC	CTGTGTCCCG	GAGGTGAAGG	CTTCAGACCT	4920
	AACCCCATCA	CAATCATTTT	AGAAGACATT	GACGAATGCC	AGGAGTTACC	AGGTCTCTGC	4980
	CAGGGTGGAA	ACTGCATCAA	CACTTTTGGG	AGCTTCCAGT	GTGAGTGCCC	ACAAGGCTAC	5040
	TACCTCAGCG	AGGATACCCG	CATCTGTGAG	GATATTGATG	AGTGTTTTGC	ACATCCTGGT	5100
	GTGTGTGGGC	CTGGGACCTG	CTATAACACC	CTGGGAAATT	ACACCTGCAT	TTGCCACCT	5160
	GAGTACATGC	AGGTCAATGG	AGGCCACAAC	TGCATGGACA	TGAGAAAAAG	CTTTTGCTAC	5220
15	CGAAGCTATA	ATGGAACCAC	TTGTGAGAAT	GAGTTGCCTT	TCAATGTGAC	AAAAAGGATG	5280
	TGCTGCTGCA	CATATAATGT	GGGCAAAGCT	GGGAACAAAC	CTTGTGAACC	ATGCCCAACT	5340
	CCAGGAACAG	CTGACTTTAA	AACCATATGT	GGAAATATTC	CTGGATTAC	CTTTGACATT	5400
	CACACAGGAA	AAGCTGTTGA	CATTGATGAA	TGTAAAGAGA	TTCCAGGCAT	TTGTGCAAAT	5460
	GGTGTGTGCA	TTAACCAGAT	TGGCAGTTTC	CGCTGTGAAT	GCCCTACAGG	ATTAGTTTAC	5520
20	AATGACCTGC	TGTTGGTTTG	TGAAGATATA	GATGAGTGCA	GCAATGGTGA	TAATCTCTGC	5580
	CAGCGGAATG	CAGACTGCAT	CAATAGTCTT	GGTAGTTACC	GCTGTGAATG	TGCCGCGGGT	5640
	TTCAAACCTT	CACCCAATGG	GGCCTGTGTA	GATCGCAATG	AATGTTTAGA	AATTCCTAAC	5700
	GTTTGCAGTC	ATGGCTTGTG	TGTTGATCTG	CAAGGAAGTT	ACCAGTGCAT	CTGCCACAAT	5760
25	GGCTTTAAGG	CTTCTCAGGA	CCAGACCATG	TGCATGGATG	TTGATGAGTG	CGAGCGGCAC	5820
	CCATGTGGAA	ATGGAACCTG	TAAAAACACC	GTTGGATCCT	ATAACTGTCT	GTGCTACCCA	5880
	GGGTTTGAAC	TCACTCATAA	TAATGATTGG	CTGGACATAG	ATGAGTGCAG	TTCTTTTTTT	5940
	GGTCAGGTGT	GCAGAAATGG	ACGTTGTTTT	AATGAAATTG	GTTCTTTCAA	GTGTCTATGT	6000
	AACGAAGGTT	ATGAACCTAC	CCCAGATGGC	AAAAACTGTA	TAGACACTAA	TGAGTGTGTC	6060
30	GCCCTTCCCG	GCTCTTGCTC	TCCTGGTACC	TGTCAGAATT	TGGAGGGATC	CTTCAGATGC	6120
	ATCTGTCCCC	CAGGGTATGA	AGTAAAAAGC	GAGAACTGCA	TTGATATAAA	TGAATGTGAT	6180
	GAAGATCCCA	ACATTTGTCT	TTTTGGTTCC	TGTACTAATA	CTCCAGGGGG	CTTCCAGTGC	6240
	CTCTGCCCCC	CTGGCTTTGT	ACTATCTGAT	AATGGACGGA	GATGCTTTGA	TACTCGCCAG	6300
	AGCTTCTGCT	TCACAAATTT	TGAAAATGGA	AAGTGTCTTG	TACCCAAAGC	TTTCAACACC	6360
35	ACAAAAGCAA	AATGCTGCTG	TAGTAAGATG	CCAGGAGAGG	GCTGGGGGGA	CCCCTGTGAG	6420
	CTGTGCCCCA	AAGACGATGA	AGTTGCATTT	CAGGATTTGT	GTCCATATGG	CCATGGAAC	6480
	GTCCCTAGTC	TTCATGATAC	ACGTGAAGAT	GTCAATGAGT	GTCTTGAGAG	CCCAGGCATT	6540
	TGTTCAAATG	GTCAATGTAT	CAACACCGAC	GGATCTTTTC	GCTGTGAATG	TCCAATGGGC	6600
	TACAACCTTG	ACTACACTGG	AGTACGCTGT	GTGGATACTG	ATGAGTGTTC	AATCGGCAAT	6660
40	CCGTGTGGAA	ATGGTACATG	CACCAATGTT	ATTGGGAGTT	TTGAATGCAA	TTGCAATGAA	6720
	GGCTTTGAGC	CAGGGCCCAT	GATGAATTGT	GAAGATATCA	ACGAATGTGC	CCAGAACCCA	6780
	CTGCTGTGTG	CTTTACGCTG	CATGAACACT	TTTGGGTCTC	ATGAATGCAC	GTGCCCGATT	6840
	GGCTATGCCC	TCAGGGAAGA	TCAAAAGATG	TGCAAAGATC	TGGATGAATG	TGCTGAAGGG	6900
	TTACACGACT	GTGAATCTAG	GGGCATGATG	TGTAAGAATC	TAATCGGCAC	CTTCATGTGC	6960
45	ATCTGCCCTC	CTGGAATGGC	CCGAAGGCCC	GATGGAGAAG	GCTGTGTAGA	TGAAAATGAA	7020
	TGCAGGACCA	AGCCAGGAAT	CTGTGAAAAA	GGACGTTGTG	TTAACATTAT	TGGAAGCTAT	7080
	AGATGTGAGT	GTAATGAAGG	ATTCCAGTCA	AGTTCTTCAG	GCACTGAATG	CCTTGACAAT	7140
	CGACAGGGTC	TCTGCTTTGC	AGAGGTACTG	CAGACAATAT	GTCAAATGGC	ATCCAGTAGT	7200
	CGCAATCTCG	TCACTAAGTC	AGAATGCTGC	TGTGATGGTG	GGCGAGGCTG	GGGCCACCAG	7260
50	TGCGAGCTTT	GCCCACTTCC	TGGAACCTGC	CAGTACAAAA	AGATATGTCC	TCATGGCCCA	7320
	GGATATACAA	CTGATGGAAG	AGATATTGAT	GAATGTAAGG	TAATGCCAAA	CCTCTGCACC	7380
	AATGGTCAGT	GCATCAATAC	CATGGGCTCA	TTCCGATGCT	TCTGCAAGGT	TGGCTACACC	7440
	ACAGACATCA	GTGGAACCTC	TTGTATAGAC	CTTGATGAAT	GCTCCCAGTC	CCCGAAACCA	7500
	TGCAACTACA	TCTGCAAGAA	CACTGAGGGG	AGTTATCAGT	GTTTATGTC	GAGGGGGTAT	7560
55	GTCTTGCAAG	AGGATGGAAA	GACATGCAAA	GACCTTGATG	AATGTCAAAC	AAAGCAGCAT	7620
	AACTGCCAGT	TCCTCTGTGT	CAACACCCTG	GGGGGGTTTA	CCTGTAAATG	TCCACCTGGT	7680
	TTACACAGC	ATCACACTGC	TTGTATCGAC	AACAACGAAT	GTGGGTCTCA	ACCTTTGCTT	7740
	TGTGGAGGAA	AGGGAATCTG	TCAAAACACT	CCAGGCAGTT	TCAGCTGTGA	ATGCCAAAGA	7800
	GGGTTCTCTC	TTGATGCCAC	CGGACTGAAC	TGTGAAGATG	TTGATGAATG	TGATGGGAAC	7860
60	CACAGGTGCC	AACACGGCTG	CCAGAATATC	CTGGGTGGCT	ACAGATGTGG	CTGCCCCCAA	7920
	GGCTACATCC	AGCACTACCA	GTGGAATCAG	TGTGTGATG	AGAATGAATG	CTCCAATCCC	7980
	AATGCCTGTG	GCTCTGCTTC	CTGCTACAAC	ACCCTGGGGA	GTTACAAGTG	CGCCTGCCCC	8040
	TCGGGGTTCT	CCTTGGACCA	GTTCTCCAGT	GCCTGCCACG	ACGTGAATGA	GTGCTCGTCC	8100
	TCCAAGAACC	CCTGCAATTA	CGGCTGCTCT	AACACGGAGG	GGGGCTACCT	CTGTGGCTGC	8160
65	CCCCCTGGGT	ATTACAGAGT	GGGACAAGGC	CACGTGTGCT	CAGGAATGGG	ATTTAACAAG	8220
	GGGCAGTACC	TGTCACTGGA	TACAGAGGTC	GATGAGGAAA	ATGCTCTGTC	CCCAGAAGCA	8280
	TGCTACGAGT	GCAAAATCAA	CGCTATGCTT	AAGAAAGACA	GCAGGCAGAA	GAGAAGTATT	8340
	CATGAACCTG	ATCCCACTGC	TGTTGAACAG	ATCAGCCTAG	AGAGTGTGGA	CATGGACAGC	8400
	CCCGTCAACA	TGAAGTTCAA	CCTCTCCCAC	CTCGGCTCTA	AGGAGCACAT	CCTGGAACCTA	8460
70	AGGCCCCGCA	TCCAGCCCCC	CAACAACACC	ATCCGTTATG	TCATCTCTCA	AGGGAACGAT	8520
	GACAGCGTCT	TCCGCATCCA	CCAAAGGAAT	GGGCTCAGCT	ACTTGACAC	GGCCAAGAAG	8580
	AAGCTCATGC	CCGGCACATA	CACACTGGAA	ATCACTAGCA	TCCCTCTCTA	CAAGAAGAAG	8640
	GAGCTTAAGA	AACCTGGAAG	GAGCAATGAG	GATGACTACC	TCCTAGGGGA	GCTTGGGGAG	8700
	GCTCTCAGAA	TGAGGCTGCA	GATTGAGCTC	TATTAACCGT	TCACAGACTT	GGGCCACGGC	8760
75	TCAAATCCTA	GCACAGCCAG	TCTGCAGAAG	CATTTGAAAA	GTCAAGGACT	AATTTTAAAG	8820
	AGGAAAAATA	ATAATAACTC	TTGTTTCTTT	CCTCCCTGTC	TTAGACTTTG	AATGTTGACC	8880
	CTCACAGGGA	GGGATAATTT	AGACTCTGGT	ATGGCCAAAG	ATTTGAGCTC	AAAGGCAACC	8940
	GTGGTTACTG	TATTTTTTAT	ATAACTTCAT	TTTAAATAT	ATTTAAAGAA	ACCTAAATGT	9000
	TCAAGATATC	AGCATATGGC	ACTAAATGCA	CAAAAATAAT	GTGAGCTTTT	TTTTTTTTTT	9060
80	CCTGTTAGCA	GTCTGTAAAC	CTTTGGGTAT	TTTGCTATAG	TTGCTAATTA	AAAAAATATA	9120
	GATGTTTATT	TATTTTAAAT	GCAGTAATAT	ATGGAGAAAT	GAACAAACTA	TGTAAACAAA	9180
	AAGGGAAACT	CACCTGTTTT	TCTTTAGATT	TATAAATTTG	AGCTATTTT	TTTAGAGGTG	9240
	CTTTTTAAAA	ATCCAATAGA	TACAAGAGAT	GTTTCCTTTG	GTTTTCTGCC	AGTCATCCAG	9300
	CTGATACACA	CCTGATCGAT	TTTAAAGAAA	GCCACACAGA	GCTGAATCGG	GCAGTGCTAA	9360
85	TCAATAATTT	AAAAGACATG	AATGTCTATTA	GATCCTTTAT	AACGTAGATC	GAAGCCAAAG	9420
	CAGCTCATT	GTGACAAACAT	TTTATATCAC	CAGACACACC	AGGCAACAGA	AGTTGAAGCA	9480
	CAACCACTGT	AGCAAAATAC	CTTGACTGCT	TGTGAGACCA	TTAGCATTGC	AGGCCAAACC	9540
	GTACTGTATT	TCCTTCTCAT	AACCTCAAGG	AACCATATGT	GCTACCCACA	ACACCTCAT	9600



5  
10  
CTTACCCAGG GTGCGCTGCG TCCTCATGGT ACTGTAGGCA GCTGAAGAAC CGCCGTTCCC 9660  
TTGAAAGGGA ACACCTGGCA TTCTGTGGTG TTTCTGTGCTG TCTTAAATAA TGGTGCATTT 9720  
ATTATGTTCA AGTTATTTCA GGATTGCCAT ATGTGCAAAC AAATCATGCA ATGCAGCCAA 9780  
GGAATATATG TTGTTGTTGT TGTTTTAAAC CCATTTTTTT TTTAGAATTT TCATTAATAC 9840  
TGTAATTATA CACCATATGC CTCATTTTAT CATAGCCTAT TGTGTATGAA AGATGTTTGT 9900  
ACAATGAATT GATGTTTAGT TTGCTTTAGT CATTAAAAA GATATTGTAC CAGGATGTGC 9960  
TATTAAGAGC ACGTATCCAT TATCTTCTC AACCCAGAA CCTGTTTCCT GGACCAAGTA 10020  
CCAAACCTCA TATGTGAAAT GGCCAAAGCA CATGCAGGCT CCTGGTTGTT CCTCTCAAAC 10080  
CTGTGCTGAC CAAAGATTAG TAACCAGTTA TACCCAGTAT TTTGAGGTTT TATTGTTTTT 10140  
TTAATAACTA AAAAAAACT CGTGCC

Seq ID NO: 459 Protein sequence  
Protein Accession #: NP\_001990.1

15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
1  
11  
21  
31  
41  
51  
MGRRLRLCLQ LYFLWLGCVV LWAQGTAGQP QPPPPKPPRP QPPPQQVRS TAGSEGGFLA 60  
PEYREGAAY ASRVRRRGQQ DVLRGPNVCS SRFHSYCCPG WKTLPGGNQC IVPICRNSCG 120  
DGFCSRPNMC TCSSGQISST CGSKSIQCS VRCMNGGTCA DDHCQCQKGY IGTTCGQPV 180  
ENGCCNGGRC IAQPCACVYG FTGFPQCRDY RTGPCFTQVN NQMCQQLTG IVCTKTLLCA 240  
TTGRAWGHP CMCAPQPPC RRGFIPNIRT GACQDVDECQ AIPGICQGGN CINTVGSFEC 300  
RCPAGHKQSE TTQKCEDIDE CSIPGICET GECSNTVGSY FCVCPRGYVT STDGSRCIDQ 360  
RTGMCFSGLV NGRCAQELPG RMTKMQCCCE PGRCWGIGTI PEACPVRGSE EYRRLCMDGL 420  
PMGGIPGSAG SRPGGTGGNG FAPSGNGNGY GPGGTGFIPI PGGNGFSPGV GGAGVAGAGG 480  
GPIITGLTIL NQIDICKHH ANLCLNGRCI PTVSSYRCEC NMGYKQDANG DCIDVDECTS 540  
NPCTNGDCVN TPGSYCKCH AGFQRTPTKQ ACIDIDEIQ NGVLCKNGRC VNSDGSFQCI 600  
CNAGFELTDD GKNVDHDEC TTTNMLNMG CINEDGSFKC ICKPGFVLAP NGRYCTDVDE 660  
CQTPGICMNG HCINSEGSFR CDCPPGLAVG MDGRVCVDTH MRSTCYGGIK KGVCVRPFP 720  
AVTKSECCCA NPDYGFGEPC QPCPAKNSAE FHGLCSSGVG ITVDGRDINE CALDPDICAN 780  
GICENLRGSY RCNCSGYEP DASGRNCIDI DECLVNRLLC DNGLCRNTPG SYSCTCPPGY 840  
VFRTEETCE DINECESNPC VNGACRNLG SFNCECSPGS KLSSTGLICI DSLKGTCLWN 900  
IQDSRCEVNI NGATLKSECC ATLGAAWGSF CERCELDTAC PRGLARIKGV TCEVDNECEV 960  
FPGVCPNGRC VNSKGSFHC CPEGLTLDGT GRVCLDIRME QCYLKWDEDE CIHPVPGKFR 1020  
MDACCAVGA AWGTECEEC KPGTKEYETL CPRGAGFANR GDVLTGRPFY KDINECKAFP 1080  
GMCTYGKCRN TIGSFKCRN SGFALDMEER NCTDIDECRI SPDLCSGIC VNTPGSFECE 1140  
CFEGYESGFM MMKNCMDIDG CERNPLLCRG GTCVNTGSGF QDCDPLGHEL SPSREDCVDI 1200  
NECSLSDNLC RNGKCVNMIG TYQSCSNPGY QATPDRQGCT DIDECEMIMNG GCDTQCTNSE 1260  
GSYECSCSEG YALMPDGRSC ADIDECENNY DICDGGCTN IPGEYRCLCY DGFMASSDMK 1320  
TCIDVNECDL NSNICMFGE CENKGSFICH CQLGYSVKKG TTGCTDVDEC EIGAHNCDMH 1380  
ASCLNIPGSF KCSCREGWIG NGIKCIDLDE CSNGTHQCSI NAQCVNTPGS YRCACSEGFT 1440  
GDGFTCSVD ECAENINLCE NGQCLNVPGA YRCECEMGFT PASDSRSCQD IDECSFQNIC 1500  
VSGTCNNLPG MFHCICDDGY ELDRTGGNCT DIDECADPIN CVNGLCVNTP GRYECNCPD 1560  
FQLNPTGVGC VDNRVGNCYL KFGPRGDGSL SCNTEIGVGV SRSSCCCSLG KAWGNPCETC 1620  
PVPNSTEYIT LCPGGEGRFP NPITILEDI DECQLPGLC QGGNCINTFG SFQCECPQGY 1680  
YLSEDRICE DIDECEFAHP VCGPCTYNT LGNYTCICPP EYMQVNGGHN CMDMRKSFCY 1740  
RSYNGTTCEN ELFPNVTKRM CCCTYNVGA GNKPCEPCPT PGTADFKTIC GNIPGFTFDI 1800  
HTGKAVDIDE CKEIPGICAN GVCINQIGSF RCECPTGFSY NDLLVCEDI DECSNGDNLC 1860  
QRNADCINSP GSYRCECAAG FKLSPNGACV DRNECLEIPN VCSHGLCVDL QGSYQCICHN 1920  
GFKASQDQTM CMDVDECERH PCGNGTCKNT VGSYNCLCYP GFELTHNNDL LDIDECSFF 1980  
GQVCRNGRCF NEIGSFKCLC NEGYELTPDG KNCIDTNECV ALPGSCSPGT CQNLGFSFR 2040  
ICPPGYEVKS ENCIDINECD EDPNICLFPS CTNTPGGFQC LCPPGFVLS DNGRRCFDTRQ 2100  
SFCFTNFENG KCSVPKAFNT TKAKCCCSKM PGEWGDPCE LCPKDDEVAF QDLCPYGHGT 2160  
VPSLHDTRED VNECLESPGI CSNGQCINTD GSFRCECPMG YNLDTGVRC VDTDECSIGN 2220  
PCGNGTCTNV IGSFECNCNE GFEPGPMNC EDINECAQNP LLCALRCMNT FGSYECTCPI 2280  
GYALREDQKM CKDLDECAEG LHDCESRGMM CKNLIGTFMC ICPPGMARRP DGECCVDENE 2340  
CRTKPGICEN GRCVNIIGSY RCECNEGFQS SSGTECLDN RQGLCFAEVL QTCIQMASSS 2400  
RNLVTKSECC CDGGRGWGHQ CELCPLPSTA GYKIKCPHGP GYTTDGRDID ECKVMPNLCT 2460  
NGQCINTMGS FRCFCKVGYT TDISGTSCID LDECSQSPKP CNYICKNTEG SYQCSCPRGY 2520  
VLQEDGKTCK DLDECQTKQH NCQFLCVNTL GGFTCKCPPG FTQHHTACID NNECGSQPLL 2580  
CGGKGICQNT PGSFSECEQR GFSLDATGLN CEDVDECDGN HRCQHGQNI LGGYRCGCPQ 2640  
GYIQHYQWNQ CVDENECSNP NACGSASCYN TLGSYKACAP SGFSFDQFSS ACHDVNECSS 2700  
SKNPNYGCN NTEGGYLCGC PPGYRVGQG HCVSGMGFNK GQYLSLDTEV DEENALSPEA 2760  
CYECKINGYP KKDSRQKRSI HEPDPTAVEQ ISLESVMDMS PVMNMFNLSH LGSKEHILEL 2820  
RPAIQPLNNH IRYVISQND DSVFRIHQRN GLSYLHTAKK KLMPGTYTLE ITSIPLYKKK 2880  
ELKKLEESNE DDYLLGELGE ALRMRLQIQL Y

Seq ID NO: 460 DNA sequence  
Nucleic Acid Accession #: NM\_013372.1  
Coding sequence: 63..617

70  
75  
80  
85  
1  
11  
21  
31  
41  
51  
GCGGCCGCAC TCAGCGCCAC GCGTCGAAAG CGCAGGCCCC GAGGACCCGC CGCACTGACA 60  
GTATGAGCCG CACAGCCTAC ACGGTGGGAG CCCTGCTTCT CCTCTTGGGG ACCCTGCTGC 120  
CGGCTGCTGA AGGGAAGAAAG AAAGGTGCCA AAGGTGCCAT CCCCCGCCA GACAAGGCC 180  
AGCACAATGA CTCAGAGCAG ACTCAGTCGC CCCAGGACC TGGCTCCAGG AACCGGGGGC 240  
GGGGCCAAGG GCGGGGCACT GCCATGCCCG GGGAGGAGGT GCTGGAGTCC AGCCAAGAGG 300  
CCCTGCATGT GACGGAGCGC AAATACCTGA AGCGAGACTG GTGCAAAACC CAGCCGCTTA 360  
AGCAGACCAT CCACGAGGAA GGCTGCAACA GTCGCACCAT CATCAACCGC TTCTGTTACG 420  
GCCAGTGCAA CTCTTTCTAC ATCCCCAGGC ACATCCGGAA GGAGGAAGGT TCCTTTTCA 480  
CCTGCTCCTT CTGCAAGCCC AAGAAATTC CTACCATGAT GGTCAACTC AACTGCCCTG 540  
AATACAGCC ACCTACCAAG AAGAAGAGAG TCACACGTGT GAAGCAGTGT CGTTGCATAT 600  
CCATCGATTT GGATTAAGCC AAATCCAGGT GCACCCAGCA TGTCTTAGGA ATGCAGCCCC 660  
AGGAAGTCCC AGACCTAAAA CAACCAGATT CTTACTTGGC TTAACCTAG AGGCCAGAAG 720  
AACCCCCAGC TGCCTCCTGG CAGGAGCCTG CTTGTGCGTA GTTCGTGTGC ATGAGTGTGG 780  
ATGGGTGCTT GTGGGTGTTT TTAGACACCA GAGAAAACAC AGTCTCTGCT AGAGAGCACT 840  
CCCTATTTTG TAAACATATC TGCTTTAATG GGGATGTACC AGAAACCCAC CTCACCCCGG 900



CTCACATCTA AAGGGGCGGG GCCGTGGTCT GGTTCCTGACT TTGTGTTTTT GTGCCCTCCT 960  
GGGGACCAGA ATCTCCTTTC GGAATGAATG TTCATGGGAG AGGCTCCTCT GAGGGCAAGA 1020  
GACCTGTTTT AGTGCTGCAT TCGACATGGA AAAGTCCTTT TAACCTGTGC TTGCATCCTC 1080  
CTTTCCTCCT CCTCCTCACA ATCCATCTCT TCTTAAGTTG ATAGTGAATA TGTCACTCTA 1140  
5 ATCTCTTGTT TGCCAAGGTT CCTAAATTAA TTCACTTAAC CATGATGCAA ATGTTTTTCA 1200  
TTTTGTGAAG ACCCTCCAGA CTCTGGGAGA GGCTGGTGTG GGCAAGGACA AGCAGGATAG 1260  
TGGAGTGAGA AAGGGAGGGT GGAGGGTGAG GCCAAATCAG GTCCAGCAAA AGTCAGTAGG 1320  
GACATTGCAG AAGCTTGAAA GGCCAATACC AGAACACAGG CTGATGCTTC TGAGAAAGTC 1380  
TTTTCTAGT ATTTAACAGA ACCCAAGTGA ACAGAGGAGA AATGAGATTG CCAGAAAGTG 1440  
10 ATTAACCTTG GCCGTGCAA TCTGCTCAA CCTAACACCA AACTGAAAAC ATAAATACTG 1500  
ACCACTCCTA TGTTCGGACC CAAGCAAGTT AGCTAAACCA AACCAACTCC TCTGCTTTGT 1560  
CCCTCAGGTG GAAAAGAGAG GTAGTTTGA ACTCTCTGCA TAGGGGTGGG AATTAATCAA 1620  
AAACCKCAGA GGCTGAAATT CCTAATACCT TTCCTTTATC GTGGTTATAG TCAGCTCATT 1680  
TCCATTCCAC TATTTCCCAT AATGCTTCTG AGAGCCACTA ACTTGATTGA TAAAGATCCT 1740  
15 GCCTCTGCTG AGTGTACCTG ACAGTAAGT TAAAGATGAR AGAGTTTAGG GACTACTCTG 1800  
TTTTAGCAAG ARATATTKTG GGGGTCTTTT TGTTTTAACT ATTGTCAGGA GATTGGGCTA 1860  
RAGAGAAGAC GACGAGAGTA AGGAAATAAA GGGRATTGCC TCTGGCTAGA GAGTAAGTTA 1920  
GGTGTTAATA CCTGGTAGAA ATGTAAGGGA TATGACCTCC CTTTCTTTAT GTGCTCACTG 1980  
AGGATCTGAG GGGACCTGT TAGGAGAGCA TAGCATCATG ATGTATTAGC TGTTCATCTG 2040  
20 CTACTGGTTG GATGGACATA ACTATTGTA CTATTAGTA TTTACTGGTA GGCAGTGTCC 2100  
TCTGATTAAA CTTGGCCTAC TGGCAATGGC TACTTAGGAT TGATCTAAGG GCCAAAGTGC 2160  
AGGGTGGGTG AACTTTATTG TACTTTGGAT TTGGTTAACC TGTTTTCTTC AAGCCTGAGG 2220  
TTTTATATAC AAACCTCCTG AATACTCTTT TTGCCTTGTA TCTTCTCAGC CTCCTAGCCA 2280  
AGTCCATATG AATATGAAA ACAAACTCTG CAGACTTGAG ATTCAGTTGC CGATCAAGGC 2340  
25 TCTGGCATTC AGAGAACCCT TGCAACTCGA GAAGCTGTTT TTATTTCTGT TTTGTTTTGA 2400  
TCCAGTGCTC TCCCATCTAA CAACTAAACA GGAGCCATTT CAAGGCGGGA GATATTTTAA 2460  
ACACCCAAAA TGTTGGGTCT GATTTTCAAA CTTTAAACT CACTACTGAT GATTCTCACG 2520  
CTAGGCGAAT TTGTCCAAAC ACATAGTGTG TGTGTTTTGT ATACACTGTA TGACCCACAC 2580  
30 CCAAATCTTT GTATTGTCCA CATTCTCCA CAATAAGCA CAGAGTGGAT TTAATTAAGC 2640  
ACACAAATGC TAAGGCAGAA TTTTGAGGGT GGGAGAGAAG AAAAGGGAAA GAAGCTGAAA 2700  
ATGTAAAACC ACACCAGGGA GGAAAAATGA CATTCAAGAC CAGCAACAC TGAATTTCTC 2760  
TTGTTGTTTT AACTCTGCCA CAAGAATGCA ATTTCTGTTA TGGAGATGAC TTAAGTTGGC 2820  
AGCAGTAATC TTCTTTTAGG AGCTTGTACC ACAGTCTTGC ACATAAGTGC AGATTTGGCT 2880  
35 CAAGTAAAGA GAATTTCTC AACACTAAGT TCACTGGGAT AATCAGCAGC GTAACCTACC 2940  
TAAAAGCATA TCACTAGCCA AAGAGGGAAA TATCTGTTCT TCTTACTGTG CCTATATTAA 3000  
GACTAGTACA AATGTGGTGT GTCTTCCAAC TTTTATTGAA AATGCCATAT CTATACCATA 3060  
TTTTATTCTGA GTCATGATG ATGTAATGAT ATATTTTTTC ATTATTATAG TAGAATATTT 3120  
TTATGGCAAG ATATTGTGG TCTTGATCAT ACCTATTAAA ATAATGCCAA ACACCAAATA 3180  
40 TGAATTTTAT GATGTACACT TTGTGCTTGG CATTAAAAGA AAAAAACACA CATCTGGAA 3240  
GTCTGTAAGT TGTTTTTGT TACTGTAGGT CTTCAAAGT AAGAGTGTA GTGAAAATC 3300  
TGGAGGAGAG GATAATTCC ACTGTGTGGA ATGTGAATAG TTAATGAAA AGTTATGGTT 3360  
ATTTAATGTA ATTATTACTT CAAATCCTTT GGTCACTGTG ATTTCAAGCA TGTTTTCTTT 3420  
TTCTCCTTTA TATGACTTTC TCTGAGTTGG GCAAAGAAGA AGCTGACACA CCGTATGTTG 3480  
45 TTAGAGTCTT TTATCTGGTC AGGGGAAAACA AAATCTTGAC CCAGCTGAAC ATGTCTTCCT 3540  
GAGTCAGTGC CTGAATCTTT ATTTTTTAAA TTGAATGTTT CTTAAAGGTT AACATTTCTA 3600  
AAGCAATATT AAGAAAGACT TTAATGTGTA TTTTGGAAGA CTTACGATGC ATGTATACAA 3660  
ACGAATAGCA GATAATGATG ACTAGTTTAC ACATAAAGT CTTTAAAGGA GAAAATCTAA 3720  
AATGAAAAGT GGATAAACAG AACATTTATA AGTGATCAGT TAATGCCTAA GAGTGAAAGT 3780  
AGTTCTATTG ACATTCCTCA AGATATTTAA TATCAACTGC ATTATGTATT ATGTCTGCTT 3840  
50 AAATCATTTA AAAACGGCAA AGAATTATAT AGACTATGAG GTACCTTGCT GTGTAGGAGG 3900  
ATGAAAGGGG AGTTGATAGT CTCATAAAAC TAATTTGGCT TCAAGTTTCA TGAATCTGTA 3960  
ACTAGAATTT AATTTTCACC CCAATAATG TCTATATAGC CTTTGCTAAA GAGCAACTAA 4020  
TAAATTAAAC CTATCTTTC AAAAAAAAA

Seq ID NO: 461 Protein sequence  
Protein Accession #: NP\_037504.1

1 11 21 31 41 51  
| | | | |  
60 MSRTAYTVGA LLLLLGTLTP AAEKKKKSQ GAIPPPDKAQ HNDSEQTQSP QPGRSRNRGR 60  
GQGRGTAMPG EEVLESSQEA LHVTERKYLK RDWCKTQPLK QTIHEEGCNS RTIINRFCYG 120  
QCNSFYIPRH IRKEEGSFQS CSFCKPKKFT TMMVTNLNCP LQPPTKKKRV TRVKQCRGIS 180  
IDL

Seq ID NO: 462 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..2733

1 11 21 31 41 51  
| | | | |  
70 ATGAAAGTTG GAGTGCTGTG GCTCATTTCT TTCTTCACCT TCACTGACGG CCACGGTGGC 60  
TTCTGGGGA AAAATGATGG CATCAAAACA AAAAAAGAAC TCATTGTGAA TAAGAAAAA 120  
CATCTAGGCC CAGTCGAAGA ATATCAGCTG CTGCTTCAGG TGACCTATAG AGATTCCAAG 180  
75 GAGAAAAGAG ATTTGAGAAA TTTTCTGAAG CTCTTGAAGC CTCCATTATT ATGGTCACAT 240  
GGGCTAATTA GAATTATCAG AGCAAAGGCT ACCACAGACT GCAACAGCCT GAATGGAGTC 300  
CTGCAGTGTA CCTGTGAAGA CAGCTACACC TGGTTTCCTC CCTCATGCC TGTATCCCAG 360  
AAGTGTACCT TACACACGGC TGGAGCACTC CCAAGCTGTG AATGTCATCT CAACAACCTC 420  
AGCCAGAGTG TCAATTTCTG TGAGAGAACA AAGATTTGGG GCACTTTCAA AATTAATGAA 480  
AGGTTTACAA ATGACCTTTT GAATTCATCT TCTGCTATAT ACTCCAAATA TGCAATGGA 540  
80 ATTGAAATTC AACTTAAAA AGCATATGAA AGAATTCAAG GTTTTGAGTC GGTTCAGGTC 600  
ACCAATTTT GAAATGGAAG CATCGTTGCT GGGTATGAAG TTGTTGGCTC CAGCAGTGCA 660  
TCTGAACTGC TGTCAGCCAT TGAACATGTT GCCGAGAAGG CTAAGACAGC CCTTCACAAG 720  
CTGTTTCCAT TAGAAGACGG CTCTTTCAGA GTGTTCCGAA AAGCCAGTG TAATGACATT 780  
85 GTCTTTGGAT TTGGGTCCAA TGATGATGAA TATACCTGCT CCTGCAGCAG TGGCTACAGG 840  
GGAAACATCA CAGCCAAGTG TGAGTCTCT GGGTGGCAGG TCATCAGGGA GACTTGTGTG 900  
CTCTCTCTGC TTGAAGAACT GAACAAGAA TTCAGTATGA TTGTAGGCAA TGCCACTGAG 960  
GCAGCTGTGT CATCTTCTGT GCAAAATCTT TCTGTCTATCA TTCGGCAAAA CCCATCAACC 1020



	ACAGTGGGGA	ATCTGGCTTC	GGTGGTGTCTG	ATTCTGAGCA	ATATTTTCATC	TCTGTCACTG	1080
	GCCAGCCATT	TCAGGGTGTC	CAATTCAACA	ATGGAGGATG	TCATCAGTAT	AGCTGACAAT	1140
	ATCCTTAATT	CAGCCTCAGT	AACCAACTGG	ACAGTCTTAC	TGCGGGAAGA	AAAGTATGCC	1200
5	AGCTCACGGT	TACTAGAGAC	ATTAGAAAAC	ATCAGCACTC	TGGTGCCTCC	GACAGCTCTT	1260
	CCTCTGAATT	TTTCTCGGAA	ATTCAATTGAC	TGGAAAGGGA	TTCCAGTGAA	CAAAAGCCAA	1320
	CTCAAAAGGG	GTTACAGCTA	TCAGATTAAA	ATGTGTCCCC	AAAATACATC	TATTTCCATC	1380
	AGAGGCCGTG	TGTTAATTGG	GTCAAGACAA	TTCCAGAGAT	CCCTTCCAGA	AACTATTATC	1440
	AGCATGGCCT	CGTTGACTCT	GGGGAACATT	CTACCCGTTT	CCAAAAATGG	AAATGCTCAG	1500
10	GTCAATGGAC	CTGTGATATC	CACGGTTATT	CAAACTATT	CCATAAATGA	AGTTTTCTTA	1560
	TTTTTTTCCA	AGATAGAGTC	AAACCTGAGC	CAGCCTCATT	GTGTGTTTGT	GGATTTTCAGT	1620
	CATTTGTCAGT	GGAACGATGC	AGGCTGCCAC	CTAGTGAATG	AAACTCAAGA	CATCGTGACG	1680
	TGCCAATGTA	CTCACTTGAC	CTCCTTCTCC	ATATTGATGT	CACCTTTTGT	CCCCTCTACA	1740
	ATCTTCCCCG	TTGTAAAATG	GATCACCTAT	GTGGGACTGG	GTATCTCCAT	TGGAAGTCTC	1800
15	ATTTTATGCC	TGATCATCGA	GGCTTGTGTT	TGGAAGCAGA	TTAAAAAAG	CCAAACCTCT	1860
	CACACACGTC	GTATTTGCAT	GGTGAACATA	GCCCTGTCCC	TCTTGATTGC	TGATGTCTGG	1920
	TTTATTGTTG	GTGCCACAGT	GGACACCACG	GTGAACCCTT	CTGGAGTCTG	CACAGCTGCT	1980
	GTGTTCTTTA	CACACTTCTT	CTACCTCTCT	TTGTTCTTCT	GGATGCTCAT	GCTTGGCATC	2040
	CTGCTGGCTT	ACCGGATCAT	CCTCGTGTTC	CATCACATGG	CCCAGCATT	GATGATGGCT	2100
20	GTTGGATTGT	GCAATGGTTA	TGGGTGCCCT	CTCATTATAT	CTGTCAATAC	CATTGCTGTC	2160
	ACGCAACCTA	GCCTAACCTA	CAAAAGGAAA	GATGTGTGTT	GGCTTAACTG	GTCCAATGGA	2220
	AGCAAACCAC	TCCTGGCTTT	TGTTGTCCCT	GCACTGGCTA	TTGTGGCTGT	GAACCTCGTT	2280
	GTGGTGCTGC	TAGTTCTCAC	AAAGCTCTGG	AGGCCGACTG	TTGGGGAAAG	ACTGAGTCGG	2340
	GATGACAAGG	CCACCATCAT	CCGCGTGGGG	AAGAGCCTCC	TCATTCTGAC	CCCTCTGCTA	2400
25	GGGCTCACCT	GGGGCTTTGG	AATAGGAACA	ATAGTGGACA	GCCAGAATCT	GGCTTGGCAT	2460
	GTTATTTTGG	CTTTACTCAA	TGCATTCCAG	GGATTTTTTA	TCTTATGCTT	TGGAATACTC	2520
	TTGGACAGTA	AGCTGCGACA	ACTTCTGTTC	AACAAGTTGT	CTGCCTTAAG	TTCTTGGAAG	2580
	CAAAACAGAA	AGCAAACTC	ATCAGATTTA	TCTGCCAAAC	CCAAATTCTC	AAAGCCTTTC	2640
	AACCCACTGC	AAAACAAAGG	CCATTATGCA	TTTTCTCATA	CTGGAGATTC	CTCCGACAAC	2700
30	ATCATGCTAA	CTCAGTTTGT	CTCAATGAA	TAA			

Seq ID NO: 463 Protein sequence  
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
35							
	MKVGVWLWIS	FFFTDGHGG	FLGKNDGIKT	KKELIVNKKK	HLGPVEEYQL	LLQVTYRDSK	60
	EKRDLRNFLK	LLKPPLWWSH	GLIRIRAKA	TTDCNSLNGV	LQCTCEDSYT	WFPPSCLDPQ	120
	NCYLHTAGAL	PSCEHLNHL	SQSVNFCERT	KIWGTFKINE	RFTNDLLNSS	SAIYSKYANG	180
40	IEIQLKKAYE	RIQGFESVQV	TQFRNGSIVA	GYEVVSSSA	SELLSAIEHV	AEKAKTALHK	240
	LFPLEDGSFR	VFGKAQCNDI	VFGFGSKDDE	YTLPCSSGYR	GNITAKCESS	GWQVIRETCV	300
	LSLLEELNKN	FSMIVGNATE	AAVSSFVQNL	SVIIRQNPST	TVGNLASVVS	ILSNISLSL	360
	ASHFRVSNST	MEDVISIADN	ILNSASVTNW	TVLLREEKYA	SSRLLETLEN	ISTLVPPTAL	420
	PLNFSRKPID	WKGIPVNSQ	LKRGYSYQIK	MCPQNTSIPI	RGRVLIGSDQ	FQRSLPETII	480
45	SMASLTGLNI	LPVSKNGNAQ	VNGPVISTVI	QNSINEVFL	FFSKIESNLS	QPHCVFWDFFS	540
	HLQWNDAGCH	LVNETQDIVT	CQCTHLTSFS	ILMSFPVPST	IFPVVKWITY	VGLGISIGSL	600
	ILCLIIIEALF	WKQIKKSQTS	HTRRICMVNI	ALSLLIADVW	FIVGATVDTT	VNPSGVCTAA	660
	VFFTHFFYLS	LFFWMLMLGI	LLAYRIILVF	HHMAQHLMA	VGFCCLGYGCP	LIISVITIAV	720
	TQPSNTYKRK	DVCWLNWSNG	SKPLLAFFVP	ALAIIVAVNFV	VVLLVLTKLW	RPTVGERLSR	780
50	DDKATIIIRVG	KSLILTLPL	GLTWGFGIGT	IVDSQNLAWH	VIFALLNAFQ	GFFILCFGIL	840
	LDSKLRQLLF	NKLSALSSWK	QTEKQNSSDL	SAKPKFSPKF	NPLQNKGHYA	FSHTGDSSDN	900
	IMLTQFVSNE						

Seq ID NO: 464 DNA sequence  
Nucleic Acid Accession #: AB035089.1  
Coding sequence: 9845..10219

	1	11	21	31	41	51	
60							
	GGGCATGCAG	CCATCGGGGA	AAATCCATAG	TGCAGATAAA	GCAAGGAGGA	AGAAGAAGGA	60
	CAGTTCTAGT	AAAAGGGAGA	ACATCAATAT	AGGATGTTTT	TTAGCAATAG	AAAAAGAAGG	120
	CCAAGAGGAA	TTAGGGAGAG	AGTTATAAGA	GATCAGCAAG	GGGACAGGGT	TAGATTTGGT	180
	TTGGTTTGAA	AGCATACAGT	AAATATGATG	TCTGTCCCTG	GCAGTGTGG	CAGAGTAGGA	240
	AGGAGGAAGG	GAGGCAAGAG	ATAATATCAT	TTTCTCTGTG	CTCCAACGT	ACTTACATAT	300
65	GAGACTATTT	CCCTCTCTGC	TTTTCAAACC	TTACTGGAGT	TGTTTTCCCT	CATGAAAACC	360
	AAGAAAGGAA	AGCTAGTTAG	TCTTGTCTG	AGGTTGTTCA	ATGTATACAT	ATCTATATCT	420
	GTAGACAGAA	TCCTTGGGAA	TACAGTAATT	GACATATATT	CTGTTATTTG	ATGCTTGAAA	480
	AATCTCCTCC	ACTAACCAGT	TTCCCTATAG	ATTGCCACAA	GCACATAATA	AGAAACAATA	540
	AATAAAATGT	TCTCTTGACT	TTGTTACTTA	ACAATGCTGA	GAAAACCTTA	CAGCCTTCAT	600
	AAGGAAGTGA	GGTCCAGGAA	AATCTAGGAG	ATATTCTTAA	ACCAATCTAT	AAAGGCATTA	660
70	GTAATGCAG	GATATTTCCT	GAAAGTGTAA	TTTCCCATTG	AGGATTTGTT	TTTAATTTCT	720
	GGATTCTTGG	AGCCAATGAA	GTTGGTGTAT	GTTTATGAAA	TATCAAGAGA	CATAAGTTGG	780
	CAAGTGTTCA	TATGCAAAAA	CTTCTTGGA	TTTCTGAGTT	CTCTGTGGCA	ATATATGACA	840
	TCAGGATATG	TCCAGTCTCA	CACACCAGGA	TATGTCCTTT	CTAGCCTGTC	TATCACATGC	900
75	TAGGAGAACT	ATTTAGGAAC	AGAAAAAAT	GCCTGAAATG	ATTTCTCATT	TGAACTCATC	960
	CAAGCTTTCT	CTAAATTTAA	GCAAACTCCT	AGTCATTTTC	AGTTAGTACC	TTTCTTAAAG	1020
	TTCAACCTTC	AGGGCAAAACC	TCCGTGCTCT	AGCAGTTTAG	CCATAGTCTG	AAATTCTCTT	1080
	CCATAGATTG	GTCCCTGTGA	ACCCCGGTTT	GTCTCAGCTT	GTTATCCTGT	TTTTTTCTTC	1140
	CCTCCATTCC	CAGGATGAGC	TTGTTGCTTC	TGTCTATAGA	GACATTAGAT	TCCTTTCTTT	1200
80	TGGTACCCGA	GTAATCCAT	CCTACTCAA	TAGAGGAAGG	TCCATTTTGT	TCTTATAGCG	1260
	CTGGATGCAG	ACTCAGCTGA	GAAGACCATT	ATTCATTTTT	GGAATCTTTT	ATCTCAGATA	1320
	TTTCTCTTTC	TTTCTTTTTT	TTCTATCTTT	GGATTTTTAG	TCCATCAACG	CCCCATTAGT	1380
	CTATTCCCCG	ACTTCAATCA	GGGAACCTAT	ACCTCTTAAA	CTCATTCAGA	GACTCAAAAC	1440
	ATATATATTG	ATACAGGAGA	CCTAAGAGAA	GCATGTCTTG	GGGGTTGAGG	AAACAGGCAG	1500
85	GTGAGAAATT	TCCAGATTGG	AAACACAGCT	TCCTTTCTCC	CATCCAGCCC	CTACTTTCAG	1560
	CCTATGTGTT	TCTGGCACCT	TGTTGTAGAT	AAATCTCCCT	TGACTTTGTG	ATGTGCTGAG	1620
	AAAACAAACT	CACGGCTGGT	GTTAAAAAGG	GCCCATGACA	ATACCAAGTG	TTGGGGAGAA	1680
	TGTGGAGAAA	TCAGAACTCT	ATTCACGGTC	GTTTGGAAATG	CACACTTGTG	CAGAATTCTA	1740



	TGGAGAAGAG	TCTGGCATT	CCTCAAAATG	TTAACCTGGA	TTTACCATAT	GACCCAGCGA	1800
	TTTCATTCAT	AGGTTTATAC	TCAAAAGAAA	TGAAGAAATA	TGCCATGCAA	AAAAATGTAC	1860
	ATGAAAGGTC	ACAACATCAT	TATTCATAAT	AGTAAAAGGA	TGGAAACAAC	ACAAATGTCC	1920
5	ATCAACTTAT	GATTAAAGAA	AATCTGGTCT	ATTCATAGAA	TGGAATATTA	TTCCGACCACA	1980
	AAAAGGAATG	ATGTACTGAT	CCATGCAATG	ATGTGGACAA	ACCATGAAAA	TAACACTAGA	2040
	TTAAAGAAGC	CAGTCACAAA	AGGACTTACT	GTATGATTCC	ATTTACCTGA	AATGTTTGGA	2100
	ATAGGCAAAAT	CCATAGAAAC	AGGAGGTAGA	TTCTTGGTTT	CCAGGGTCTC	CAGGAAGGGA	2160
	AGAATGAAGT	ACAAGATTTT	TTTTGGAGGT	AGTGAAATTG	TTGTGGAATG	AGATCATGAT	2220
10	GATGATAGCA	CAACTTTGTG	AATATAATAA	AATCATTGAA	TTGTACAGTT	GAATTTATGG	2280
	TATATAAATT	ATATGTTAAT	AAAAAGGGGG	TCCACAAAAC	AAACAGCCCC	CCACTCTGGT	2340
	TGTCAGGGAG	ATATTGGATT	AAATGGCCTT	GGACAACAAC	CCCTCTCCCT	GGCCACAGAC	2400
	ATTCTTCAGA	TTACAAGATA	TTCCAGGGGA	AACACTGGAA	TGAGTCTGAA	GCCAGGTGCT	2460
	AAACAGAAGG	ACCATTGAGA	AATGTTGTGA	TCTTGACAGG	TCAAGCAATT	TATTTTTCGG	2520
15	CTTCATTTTT	AAATGTAAAA	TTAGAAAGCT	GCCATTTAAA	ATGGCCCGTC	TGTTTCAATT	2580
	GCTCTTCTCA	GTGTCAGCCT	GTAACTCAA	TGTTTTCATG	CTGCTGATAA		2640
	AAACATACCT	GAGAAAGGCA	AGAAAAAGAG	GTTTAATTGG	GCTTAGAGTT	CCACGTGATT	2700
	GGGGAGGCCT	CAGAATCACA	GTAGGAGGCA	AAAGTTATTC	TTACATGGTG	GCTGCAAGAG	2760
	AAGATGAGGA	AGAAGCAAAA	GAAGAAACCC	CTGATAAACC	CATCGGATCT	CCTGAGGCTT	2820
20	ATTAACTATC	ATGAGAATAG	CACAAGAAAG	ACCGGCCCCC	ATGATTCAAT	TACCTCTACC	2880
	TGGGTCCCTC	CAATAACATG	TGGAATTTCT	GGTAGATACA	ATTCAAGTTG	AGATTGGGT	2940
	GGGAACACAG	CCAAACCATA	TCACTCAGCA	AGGCAGATAA	CTTCTCACT	GAGCCTATGC	3000
	AACAGAAAAC	CATCTGGGAT	GGTTGTAAGG	GGCAGAGGAA	GTGACTGGTA	GGATCACTGC	3060
	CAAAGCTGAG	CACTCAGGAG	AAGGCAATAG	AATCCTATTC	TCCATAGTAT	GCTATAAGAT	3120
25	ACTGAAGTAC	ACTTCTTCAC	TATCTCTTTG	GACTTAGAAT	TAGCACTACA	TTCTTTGTTA	3180
	TACAGAAAAA	TTACTAAGGA	AATTCATAGG	ATGACAAAAA	CTTTCAGAAC	TGAAAAACAG	3240
	GAAATGTAAG	CTTTTTAGTT	CTTTGGTATT	CGAAGTATGC	CTAAAAGACA	ATGCAAAATC	3300
	CAAGAAAAGA	ATGGTGGGGT	TTTTGTTTGT	TTGGTTTTGT	TTTTGTTTTA	CAGCTGGAGT	3360
	AGAATACAAA	GGGATGGAGT	TGAAACAAAT	GAGAGGAAAT	TGGAATTCTA	AACTTATTCT	3420
30	CATTGGCATT	AGAAAGGCAC	CTACATGTAT	TTCCATGAG	CCGGTGAAGT	CTGACTTGCA	3480
	TTCTTATTTT	TTCCCTATAG	ATTAATAAGG	AGGTACAATG	GTAGAACTGT	AATCCTGTCC	3540
	TTTGTCTATA	ATTTTCATAT	TCATAAAGGT	GAGTGTTAGC	CCGCTTGTGA	AATCTGAAGT	3600
	TGAGTAACTT	CAAATACTAA	CCACAGAGGG	AAAGGCAGCA	AGAGGAGAGG	CATAAATTTA	3660
	GGATCTCACC	CTTCATTCCA	CAGACACACA	CAGCCTCTCT	GCCCACCTCT	GCTTCTCTTA	3720
35	GGAACACAGG	TAAGAGCTTC	AAGCCTCTCC	AGCTTAATAA	CATGAATTAT	TTTTTGAGAT	3780
	AATAATGATA	CTGTGTTCTA	TATCATGCAT	CTCCTGCATT	CTGTCTGATT	ATATTTTACT	3840
	TATTCGTCCA	GAGCAAAATT	AAAATACCTA	TTTCATCTGA	TTTGTCTTTT	ATCTAAATTG	3900
	CTTAGTTCCA	AGTAAACCAA	GGCACTTTTA	GGAACACAGA	GGGAGAGTGC	CTTGCAGCCA	3960
	GAGAGTCTTG	AAGGAGATGT	CAGGGACGCA	TCTTAACAGC	TGGTTGGATG	TGATCCACAG	4020
40	AGGTCTCCTG	TTAGCATTCA	TTGTAAAGCC	ATCCTACCTA	GCTCTAGTGT	AACCAGCAAT	4080
	GAAAGAAAGA	TAAAGAGGGT	CGATTACTTA	TTTACAATAG	TCTTTAAAAA	CGTAGTTTTG	4140
	TAAGCCTTCT	AATTAGGACA	TTAATATATT	TATTAATATG	ACATTGTAGA	AAGATTGAAG	4200
	CGTTAAAAAT	AAGAGAAAAA	CTTTAAATGT	CAAAATCTCA	CAACCCAGAT	ATATCATTTT	4260
	TTTAAGAAAA	TTGTACTACA	AAATACCATT	CCATTATTTA	AAGTCATTCT	GACAGGAATC	4320
45	TGATGCTTTT	CCAGGAGTTC	CAGATCACAT	CGAGTTCACC	ATGAATTAC	TCAGTGAAGC	4380
	CAACACCAAG	TTCATGTTTC	ATCTGTTCCA	ACAGTTCAGA	AAATCAAAAG	AGAACAACAT	4440
	CTTCTATTCC	CCTATCAGCA	TCACATCAGC	ATTAGGGATG	GTCTCTTAG	GAGCCAAAGA	4500
	CAACACTGCA	CAACAAATTA	GCAAGGTAGC	TATCAGCATC	ATTACGTTGT	CCTGTTGCAG	4560
	TTTTTCTCTG	GTTCCGTCGG	CTAGCACGCA	GATGGTAATA	GATGTGGTGG	TCTGATGGGT	4620
50	AGCACAGGGG	GCTGTGCAGG	AATTCCCATA	ACTGTGAGAC	CACTGACTTA	AACAGATCTT	4680
	TTGAGTAAAG	TTTTCTTGTC	CCGCTTCATG	TCTCTTCCAG	GTTCTTCACT	TTGATCAAGT	4740
	CACAGAGAAC	ACCACAGAAA	AAGCTGCAAC	ATATCATGTG	AGTCACAGAG	CACTCTGATT	4800
	CAGCTTTAGA	TCCCTGAACA	GGTCATAGTT	TAAACCTGGA	ACTTCACAAA	AACATAAGAA	4860
	AGGCCAGTTT	TGGGAAATTA	CTTGGACACA	AAGATTGAGA	CATACAGAGT	GGGTGGGCAT	4920
55	TTTCATGGCAC	ATAATTATTA	TTCTCATTTT	CTGCGTTACT	AAAAGACAGT	CAGCACTGTA	4980
	CCTCAGAGCA	TAGGTCTGGA	TCAGGATAGG	CTGGGTTTCT	ACTCCAGCTT	TGCTCTTCAC	5040
	AAATGATGAA	TAAGAGCAGG	ACACAACCTG	TCGGAGTCCC	AGTGACCTCA	TCCCAGAAAA	5100
	CTAAGGGTAA	GAAAAAATCT	GACTCAATAC	ATGCAAAATAC	ATGCAAAATGT	TTACAACAGT	5160
	GCCTTGCCCA	TAAAAGTCAT	AATAAATGTT	ATTATTATTA	TAAAGTAGCT	ATAATTATAC	5220
60	TAATCATAAT	AATGTGAAAA	TAATTTAATT	TTTATTGAGT	CATTAATGAG	ATTCAGAGGA	5280
	ATAAGCACAA	GTCCAAGTAT	ATTTTGGAAA	ATGATTGCTA	TGGAATATAT	TGGTTTAGAG	5340
	CCTTAATAGT	GCAAAATGCT	TTGCTGGAAG	GTAGAAAGTT	CTAGATTTAA	ACAGGCTTAG	5400
	GTTCAAAACT	TGGCACTTCT	AATTTATGTC	TCTATAAACA	GGGTTTTTTT	CCCCATTCTC	5460
	TGAGCTTCTT	TGTGTTCAIC	TGAATTGAAC	TAAAGACTTA	GAGTTACCCA	TGTAAAGTCC	5520
65	TTAGCCATGG	ACCTGGCATA	CACTCTTCTT	ACGTGCAGAG	AATGACCATC	ATGAGGAAAG	5580
	AGCCACAGAT	CAGTCAATGT	GTCTTACAAG	ATAATAGCAC	CAACAGGTAT	AACAGGGCTT	5640
	CCTGGCATAA	TCTATTTAAA	ATATCCAACC	TTCAACATAC	TCGTATCCTT	GATGACTGTT	5700
	AGAAGTGAAA	TATGGTCTCT	GCCCATAGG	AGCTGAGAGT	TAACTGGGA	AGCTAAACCT	5760
	AACCCTTTAA	ACCAACAAGG	AGAAAATCTA	CTGGTAGACA	GCGCTGCATC	TTTAGTTTCT	5820
70	AAGAGAAAAG	ATTGCAGTAC	GTTAGAGCAA	GAAGAAATTT	CTGGAAGAAG	TCAAATATAA	5880
	GGTGGATTTT	GAAGGGTATT	TGAGGTGAAA	TACACCAATT	ATCAGGGAAT	AACATCAAAG	5940
	GTCTCTCAAT	AGACTACCAG	CATTTAGGGA	CTGATCTAAC	AGACTTAGCA	TGGGTTTAGT	6000
	ATTTACATTG	ATACAGCAAT	TGAATGATCT	CCTTTTTTGA	TGTTTGAAGG	TTGATAGGTC	6060
	AGGAAATGTT	CATCACCAGT	TTCAAAAGCT	TCTGACTGAA	TTCAACAAAT	CCACTGATGC	6120
75	ATATGAGCTG	AAGATCGCCA	ACAAGCTCTT	CGGAGAAAAG	ACGTATCAAT	TTTTACAGGT	6180
	AATTTACCTT	GGCTACCCA	CATTTCAATT	GCATCCTGAT	GTCTGTGTCT	CTGAGTGGCC	6240
	AAATGGAAGA	AAGCAAGGCA	GATGAGCCTG	GCCGACCCAG	GTGGAGAGCA	TTTACTCAGA	6300
	GTGCATTAGC	TCCATTTCCA	CAACTCTCCC	CCACTGGAGT	GTCCAGAGCC	CCAACGATAC	6360
	ATCACTGAAG	TGTGATTTA	GGGATAATCT	TGTGATAAAA	GAGGAGGTTG	TGTAATAGAG	6420
80	TGAGTAAGAG	TAATAAGTAA	TAAGATACCA	TCGATAAACT	GGCACTGACT	CAGTCACATA	6480
	CGATACATCT	TGGTGGGAAA	TGTATGACTA	ATGGGATATT	ATTGGAATGG	GCAGGCTTGG	6540
	GTGAGTTTCT	GAGAATAGTT	GAGGAAGTAC	CAGGAAATAT	TGAATGCACA	GGATGAAAGA	6600
	CAAAAACAAA	GATCAGAAAC	ATCATGGTTA	AAATTACTGG	AGAGAAGTCT	GAGAAGCAAT	6660
	GAATCTCCTT	CAGGGAAAGCC	TGCTCTGCAG	TTTGCAAAACC	ACAGCCTCTT	CTGCTTCTGC	6720
85	CTTTTGCCAA	GATGATATTG	ACCTTCTAGT	ACCTCTTCTT	TGTGCCAGCC	CACATTCCCC	6780
	TTTTGCAATT	CCTACATGAC	ACCTGTATAA	AAATATCCAT	GGACAGGAGA	TACTGCATCT	6840
	ATTCAAGGTC	TGGATTGAGC	TTACTGTTGT	TACAAATAAG	TAAATTGGT	AATATATAGT	6900
	TACATAAATT	ACTCCTAATT	CCTACTTCTT	CCTTCATATC	TCAAAGGAAT	ATTTAGATGC	6960



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75

CATCAAGAAA TTTTACCAGA CCAGTGTGGA ATCTACTGAT TTTGCAAATG CTCCAGAAGA 7020  
AAGTCGAAAG AAGATTAACT CCTGGGTGGA AAGTCAAACG AATGGTAGGA GAGCCACCCA 7080  
TTATAGAAAC ACCTTTGAGA AACCTATGCC AGTGAGCCTT GTGCTTGACA CTGCATGGGG 7140  
GAACAGGTGT GGGGATTGAG ATGGGTTTGC AGGGAGGGCT GAAGAGGGCA CTCCAGATGA 7200  
AGGATTTGTC CAAATGAATA TGAAGAGAGC CTAGGGGAGC CAAGGAGGAA ATCACAGGAA 7260  
GCCAATTAGA TGGAAACACA TCTGGAGAAT TATTTGCTTA TGGCCCTGCA TGACAATAGC 7320  
TTTGTGGATC CCCTGTCTCC GCTCAGACCT ATTTTGAGAT CATATCCTTT ACTTTAAATC 7380  
AGACTCAAAT TTTTATGATG AATATTTAAT AGAAAACATT AGAAAGCGTC TCTCGTCTCC 7440  
TTTACTAATT GGGAAACAAG CAGCTCTCTG GTAAATCACC CTTTTGTCTC TGAGCTGGAG 7500  
CTGCCTGGAT CACATCTGTA GCCAATGTGT TCTGCAGGGA TTATCACAGC TCTCTTCCCC 7560  
ATCAAGGGCA AAGAGCTTGA CAAAGTCTCC ATTCTACAGA CATCTTCTT ACCTCCCACC 7620  
TCTCATTACA GGCCAACTT ACAGCAACTC AACATGAGAG TGAATAGGAA GATACCCCCG 7680  
GAAGTAGTGT CTGACAGCAC AGGACATGCG TTTTATATTA CAGAGCTCAA GTCACCTCATC 7740  
CTAAAGTGCA ATCAGGGCCT CCTTCCTCTG AATGGGGACC CCGTAGTTAA AAAAAAATAA 7800  
AAGTAGGAAG AGGAGGGAGG GAGAAAGGAA AGACACATGT TGGAAAGAGTA GACAAAATCA 7860  
GTTTATCAGT ATTCCAAATC AGATGATTGG AGACATTCAT ACACAGAGAA CGTGAACCTC 7920  
TTCTCTATCA CAAGAAGTGA TGTCTCCATC AAGGGTAACT TTATACGACT GGAGCCTTGA 7980  
AGAAAGCTGC ATCTGGTGAA CCACTGGTCA GTGAGTCTAA CAATTCAAAG ATCAAAGTCA 8040  
GTGAGTCTCA AGCAGGGATT TGGGTCAATA ATTAACGATC AGTCACGAAC ATTTGCAAAG 8100  
CATCTTCCAG ACAAGCCATT TGTAGCTTGT GTAAAGACT CTTTTATTCT TCCCTTGCA 8160  
GAAAAAATTA AAAACCTATT TCCTGATGGG ACTATTGGCA ATGATACGAC ACTGGTTCTT 8220  
GTGAACGCAA TCTATTTCAA AGGGCAGTGG GAGAATAAAT TTAATAAAGA AAACACTAAA 8280  
GAGGAAAAAT TTTGGCCAAA CAAGGTATTG TCTATATTTT ATTTATATAG TGTAAATATGT 8340  
TAATACATGG AATGTTAAAC ATTTCTGATG GAATGTAACA TGATAAGTAA AAAATAAAAA 8400  
TTGTTTATGT CTGTTATTTT GTTGTTTTAC TCTTATAACT TTATTTAGTT AGGAATACCT 8460  
GAAAAACTAT TGTCTTAACT TCATGGAATT CCTGGGTTAT TTCTTAGAAG AAGAAGGATG 8520  
TGTTGCTATC TCAATAATAT TATCTTTTGT GCTTGTGTGT TCACGTGTTA TTTGTTGGAC 8580  
ACATTGATTT ATTGCAGAAT ACATACAAAT CTGTACAGAT GATGAGGCAA TACAATTCCT 8640  
TTAATTTTGC CTTGCTGGAG GATGTACAGG CCAAGGTCTT GGAAATACCA TACAAAGGCA 8700  
AAGATCTAAG CATGATTGTG CTGCTGCCAA ATGAAATCGA TGGTCTGCAG AAGGTAAGAA 8760  
CTTGCACTCA CAACTCTTCC TTCTACTGCC GGACATTTT CCAAAGATAC CAAGTTTAAA 8820  
CAAGGTAAAA GCTTATGACC GAGTTGCCTC AAAATGATGA AAAATCTAA ATGAGGAATG 8880  
ATGACTCACC TTCATATTAC AAATATTTGA GCATAGGGCC TGACACAAAC TGAAAGCTTA 8940  
GTTTTTGTGT GTTTGTTTGT TTTTATTATT ATTATTATAA TACTTTAAGC TTTAGGGTAC 9000  
ATGTGCACAA TGTGCAGGTT AGTTACATAT GTATACATGT GCCATGCTGG TGTGCTGCAC 9060  
CCATTAACCTC ATCATTTAGC GTTAGGTATA TCTCCTAATG CTATCCCTCC CCCCTCCCCC 9120  
CACCCACAAA CAGTCTCAG AGTGTGATGT TACCTTCCTG TGTCCAAGTG TTCTCATTGT 9180  
TCAATTTCCA TCTATGATT AATTCATCT ATGGCTTAGT TAATGATTAA TTTATTAGAG 9240  
TTACATGCAT TGGATATCAA TTTGATGATA TTATTATGCA GCAATTTAAA CTGACTGGG 9300  
AGAAATATAT ACCAATGTGA GGAAAGTTTA CAAATAGGCC GAGTAGAAAA GGGAAATACAA 9360  
ATTTAGGAAT TTAGGGAATT ACAATTTAAT AATTGCAATG TGTACTAAAT AATGTATACA 9420  
GAAAAATATG ATGAGCCTAT TAAAAATTGA CACATGTAGT AGGCTGTTGG CACAAGAAAT 9480  
AGTGATACAT ACAGTTCATT GTGTACAAAA TAATGTAATC ATATTTTACA TGTGTATCAT 9540  
ACAGTTGTAT ACATACATAT GTACACATAT ACATATACGT AAAACATGA TTCTGTTTTT 9600  
ACATACATGT ATATACATAT ACACATATAA CCAATGTAT TTATATATT AGGACTATA 9660  
TTTTACCTAT TAGAATAATA ATGCTATTTA AAGTGAACCT TCTGTATTTT ACATTTATTG 9720  
CCAAAAAATC GAATCTCCAC ATAGTCAATT CATTGTAAAG GTGTATTAGA GATCGACAGT 9780  
TAGTCATATC AGTTTCTTTT TTCCATTTGT ATAGCTTGAA GAGAACTCA CTGCTGAGAA 9840  
ATTGATGGAA TGGACAAGTT TGCAGAATAT GAGAGAGACA TGTGTCGATT TACACTTACC 9900  
TCGGTTCAAA ATGGAAGAGA GCTATGACCT CAAGGACACG TTGAGAACCA TGGGAATGGT 9960  
GAATATCTTC AATGGGGATG CAGACCTCTC AGGCACTGACC TGGAGCCACG GTCTCTCAGT 10020  
ATCTAAAGTC ATCACAAAGG CCTTTGTGGA GGTCACTGAG GAGGGAGTGG AAGCTGCAGC 10080  
TGCCACCGCT GTAGTAGTAG TCGAATTATC ATCTCCTTCA ACTAATGAAG AGTTCTGTTG 10140  
TAATCACCTT TTCCTATTCT TCATAAGGCA AAATAAGACC AACAGCATCC TCTTCTATGG 10200  
CAGATTCTCA TCCCCATAGA TGCAATTAGT CTGTCACTCC ATTTAGAAAA TGTTACCTTA 10260  
GAGGTGTTCT GGTAACTGA TTGCTGGCAA CAACAGATTCT TCTTGGCTCA TATTTCTTTT 10320  
CTATCTCATC TTGATGATGA TAGTCATCAT CAAGAATTTA ATGATTAAAA TAGCATGCCT 10380  
TTCTCTCTTT CTCTTAATAA GCCACATAT AAATGTACTT TTCCTTCCAG AAAAATTTCC 10440  
CTTGAGGAAA AATGTCCAAG ATAAGATGAA TCATTTAATA CCGTGTCTTC TAAATTTGAA 10500  
ATATAATTCT GTTTCTGACC TGTTTTAAAT GAACCAAACC AAATCATACT TTCTCTTCAA 10560  
ATTTAGCAAC CTAGAAACAC ACATTTCTTT GAATTTAGGT GATACCTAAA TCCTTCTTAT 10620  
GTTTCTAAAT TTTGTGATT TATAAAACAC ATCATCAATA AAATAATGAC ATAAAAATCAT 10680  
TTTTGCTTTA CCTGTTTTCT CTCTGGAAAG GGCAAGTGTC CAGTTACACA TAGGAAAGAT 10740  
AATTTAGAGA TATATTAATC ATATATAAAG GAAATTTAAA AACAGAGTAG TTCATGATGA 10800  
GCTTGGAGTA GAAGGCATAT CCCAGAACAG GAGGAGCCTT GTAAACCACA TAGGAACTTC 10860  
CTATTTTATG CTAAAGGGAT AAGAACTCA TTACAGGCTT TGATGGTTGT TTGTCAAAGA 10920  
GGGGCATAAA ATTATCATAT CCACATCTAG AAAATACATC TCTGGCTACG CTGATATCAA 10980  
TGGATGCGAG GAAAGAACAG TGTGGTTACC ATATATAAAT TAGGAAATCA TTAGAGTATT 11040  
GGGAGTGGAA ATGGAGAGAA AGAAAGAGCC TGGGGGAATT ATTTAGGAAA TAATAGTTAC 11100  
AGAAAGACAT CTAAGTTGCT GACCTATCTG ACTGGATGGA TGGAGAATA TCTTGTCTCT 11160  
GAGAGAAAAA AAGACTTTGG GTTTAAATTT GTACTTGATG AATTAAGGTA CTTTAAATAT 11220  
TCAAAATGGAT TTGCCTGGCA GGCACCTGAA GATATTAGTC TAAATCTCAG AAACAGATAA 11280  
TGATCTGAAG CTCTAAATTT GTGATATTCA ATATAAATAC TTTAGAGTCA TTGGGATAAA 11340  
TATGGTAGTT GTAGCTAAAA GCAAAAATAA GATACTAGGG AGAAAGGATA AAGTTAGAAG 11400  
AAAGAAGAAT CTAGAATTGA CCTTGAAGTA TATCAGCATG TGTAAGATC AGGAATTGAT 11460  
CATTTTATT TTCCAGAAAG TAGCTTTTCT TAGGGTCCA TATTTACTCC CATAGATTCT 11520  
TCCC

Seq ID NO: 465 Protein sequence  
Protein Accession #: BAB21525.1

85  
1 11 21 31 41 51  
| | | | |  
MNSLSEANTK FMFDLFQQFR KSKENNIFYS PISITSALGM VLLGAKDNTA QQISKVLHFD 60  
QVTENTTEKA ATYHVDRSGN VHHQFQKLLT EFNKSTDAYE LKIANKLFG E KTYQFLQEYL 120  
DAIKKFYQTS VESTDFANAP EESRKKINSW VESQTNKIK NLFPDGTIGN DTTLVLVNAI 180  
YFKGQWENKF KKENTKEEFK WPNKNTYKSV QMMRQYNSFN FALLEDEVQAK VLEIPYKGD 240



LSMIVLLPNE IDGLQKLEEK LTAEKLMWEV SLQNMRETCV DLHLPRFKME ESYDLKDTLR 300  
 TMGMVNIFNG DADLSGMTWS HGLSVSKVLH KAFVEVTEEG VEAAAATAVV VVELSSPSTN 360  
 BEFCNHPFL FFIRQNKTN ILFYGRFSSP

5

Seq ID NO: 466 DNA sequence  
 Nucleic Acid Accession #: NM\_001910.1  
 Coding sequence: 50..1240

10	1	11	21	31	41	51	
	GGAGAGAAGA	AAGGAGGGGG	CAAGGGAGAA	GCTGCTGGTC	GGACTCACAA	TGAAAACGCT	60
	CCTTCTTTTG	CTGCTGGTGC	TCCTGGAGCT	GGGAGAGGCC	CAAGGATCCC	TTACAGGGT	120
	GGCCCTCAGG	AGGCATCCGT	CCCTCAAGAA	GAAGCTGCGG	GCACGGAGCC	AGCTCTCTGA	180
	GTTCTGGAAA	TCCCATTAAT	TGGACATGAT	CCAGTTCACC	GAGTCTGCT	CAATGGACCA	240
15	GAGTGCCAAG	GAACCCCTCA	TCAACTACTT	GGATATGGAA	TACTTCGGCA	CTATCTCCAT	300
	TGGCTCCCCA	CCACAGAACT	TCACTGTCT	CTTCGACACT	GGCTCCTCCA	ACCTCTGGGT	360
	CCCCTCTGTG	TACTGCACTA	GCCCAGCCTG	CAAGACGCAC	AGCAGGTTCC	AGCCTTCCCA	420
	GTCCAGCACA	TACAGCCAGC	CAGGTCAATC	TTTCTCCATT	CAGTATGGAA	CCGGGAGCTT	480
	GTCCGGGATC	ATTGGAGCCG	ACCAAGTCTC	TGTGGAAGGA	CTAACCGTGG	TTGGCCAGCA	540
20	GTTTGGAGAA	AGTGTACACG	AGCCAGGCCA	GACCTTTGTG	GATGCAGAGT	TTGATGGAAT	600
	TCTGGGCCTG	GGATACCCCT	CCTTGGCTGT	GGGAGGAGTG	ACTCCAGTAT	TTGACAACAT	660
	GATGGCTCAG	AACCTGGTGG	ACTTGGCGAT	GTTTTCTGTC	TACATGAGCA	GTAACCCAGA	720
	AGGTGGTGGC	GGGAGCGAGC	TGATTTTGGG	AGGCTACGAC	CACTCCCATT	TCTCTGGGAG	780
	CCTGAATTGG	GTCCCACTCA	CCAAGCAAGC	TTACTGGCAG	ATTGCACTGG	ATAACATCCA	840
25	GGTGGGAGGC	ACTGTATGT	TCTGCTCCGA	GGGCTGCCAG	GCCATGTGG	ACACAGGGAC	900
	TTCCCTCATC	ACTGGCCCTT	CCGACAAGAT	TAAGCAGCTG	CAAAACGCCA	TTGGGGCAGC	960
	CCCCGTGGAT	GGAGAATATG	CTGTGGAGTG	TGCCAACCTT	AACGTCATGC	CGGATGTCAC	1020
	CTTCACCAT	AACGGAGTCC	CCTATACCT	CAGCCCAACT	GCCTACACCC	TACTGGACTT	1080
	CGTGGATGGA	ATGCAGTTCT	GCAGCAGTGG	CTTTCAGGA	CTTGACATCC	ACCCTCCAGC	1140
30	TGGGCCCCCTC	TGGATCCTGG	GGGATGTCTT	CATTTCGACAG	TTTACTCAG	TCTTTGACCG	1200
	TGGGAATAAC	CGTGTGGGAC	TGGCCCCAGC	AGTCCCCATA	GGAGGGGCCT	TGTGTCTGTG	1260
	CCTGCCTGTC	TGACAGACCT	TGAATATGTT	AGGCTGGGGC	ATTCTTTACA	CCTACAAAAA	1320
	GTTATTTTCC	AGAGAATGTA	GCTGTTTCCA	GGGTGCAAC	TTGAATTAAG	ACCAAACAGA	1380
	ACATGAGAAT	ACACACACAC	ACACACACAC	ACACTTCACA	CATACACACC		1440
35	ACTCCACCA	CCGTCATGAT	GGAGGAATTA	CGTTATACAT	TCATATTTTG	TATTGATTTT	1500
	TGATTATGAA	AATCAAAAAT	TTTCACATTT	GATTATGAAA	ATCTCCAAAC	ATATGCACAA	1560
	GCAGAGATCA	TGGTATAATA	AATCCCTTTG	CAACTCCACT	CAGCCCTGAC	AACCCATCCA	1620
	CACACGGCCA	GGCCTGTTTA	TCTACACTGC	TGCCCCACTCC	TCTCTCCAGC	TCCACATGCT	1680
	GTACCTGGAT	CATTCTGAAG	CAAATTCGCA	GCATTACATC	ATTTTGTCCA	TAAATATTTT	1740
40	TAACATCCTT	AAATATACAA	TCGGAATTCA	AGCATCTCCC	ATTGTCCAC	AAATGTTTGG	1800
	CTGTTTTTGT	AGTTGGATTG	TTTGATTAG	GATTCAAGCA	AGGCCCATAT	ATTGCATTTA	1860
	TTTGAAATGT	CTGTAACTCT	CTTTCCATCT	ACAGAGTTTA	GCACATTTGA	ACGTTGCTGG	1920
	TTGAAATCCC	GAGGTGTCT	TTGACATGGT	TCTCTGAAC	TATCTTTCCT	ATAAAATGGT	1980
	AGTTAGATCT	GGAGGTCTGA	TTTGTGGCA	AAAATACTTC	CTAGGTGGTG	CTGGGTACTT	2040
45	CTTGTTGCAT	CCTGTCAGGA	GGCAGATAAT	GCTGGTGCCT	CTCTATTGGT	AATGTTAAGA	2100
	CTGCTGGGTG	GGTTTGGAGT	TCTTGGCTTT	AATCATTCAT	TACAAAGTTC	AGCATTTT	

Seq ID NO: 467 Protein sequence  
 Protein Accession #: NP\_001901.1

50	1	11	21	31	41	51	
	MKTLLELLLV	LLELGAEQGS	LHRVPLRRHP	SLKKKLRRAS	QLSEFWKSHN	LDMIQFTESC	60
	SMDQSAKEPL	INYLDMEYFG	TISIGSPPQN	FTVIFDTGSS	NLWVPSVYCT	SPACKTHSRF	120
55	QPSQSSTYSQ	PGQSFSIQYG	TGSLSGIIGA	DQVSVEGLTV	VGQQFGESVT	EPGQTFVDAE	180
	FDGILGLGYP	SLAVGGVTPV	FDNMMAQNLV	DLPMFSVYMS	SNPEGGAGSE	LIFGGYDHS	240
	FSGSLNWVPV	TKQAYWQIAL	DNIQVGGTVM	FCSEGCQAI	DTGTSITLGP	SDKIKQLQNA	300
	IGAAPVDGEY	AVECANLNMV	PDVFTTINGV	PYTLSPATY	LLDFVDGMQF	CSSGFQGLDI	360
60	HPPAGPLWIL	GDVFIRQFYS	VFDRGNRRVG	LAPAVP			

Seq ID NO: 468 DNA sequence  
 Nucleic Acid Accession #: NM\_018058.1  
 Coding sequence: 319..1575

65	1	11	21	31	41	51	
	TACGCGCTGC	GGGACCGGCA	GGGGAACGCC	ATCGGGGTCA	CAGCCTGCGA	CATCGACGGG	60
	GACGGCCGGG	AGGAGATCTA	CTTCCTCAAC	ACCAATAATG	CCTTCTCGGG	GGTGGCCACG	120
	TACACCGACA	AGTTGTTCAA	GTTCCGCAAT	AACCGGTGGG	AAGACATCCT	GAGCGATGAG	180
70	GTCAACGTGG	CCCGTGGTGT	GGCCAGCCTC	TTTGCCGGAC	GCTCTGTGGC	CTGTGTGGAC	240
	AGAAAGGGCT	CTGGACGCTA	CTCTATCTAG	ATTGCCAATT	ACGCTACGG	TAATGTGGGC	300
	CCTGATGCCC	TCATTGAAAT	GGACCCCTGAG	GCCAGTGACC	TCTCCCGGGG	CATTCTGGCG	360
	CTCAGAGATG	TGGCTGCTGA	GGCTGGGGTC	AGCAAATATA	CAGGGGGCCG	AGGCGTCAGC	420
	GTGGGCCCCA	TCCTCAGCAG	CAGTGCCTCG	GATATCTTCT	GCGACAATGA	GAATGGGCCT	480
75	AACTTCCTTT	TCCACAACCG	GGGCGATGGC	ACCTTTGTGG	ACGCTGCGGC	CAGTGTGGT	540
	GTGGACGACC	CCCACCAGCA	TGGGCGAGGT	GTGCGCCCTG	CTGACTTCAA	CCGTGATGGC	600
	AAAGTGGACA	TCGTCTATGG	CAACTGGAAT	GGCCCCCACC	GCCTCTATCT	GCAAATGAGC	660
	ACCCATGGGA	AGGTCCGCTT	CCGGGACATC	GCCTCACCCA	AGTTCTCCAT	GCCCTCCCCT	720
	GTCCGCACGG	TCATCACCGC	CGACTTTGAC	AATGACCAGG	AGCTGGAGAT	CTTCTTCAAC	780
80	AACATTGCCT	ACCGCAGCTC	CTCAGCCAAC	CGCCTCTTCC	GCGTCATCCG	TAGAGAGCAC	840
	GGAGACCCCC	TCATCGAGGA	GCTCAATCCC	GGCGACGCC	TGGAGCCTGA	GGGCCGGGGC	900
	ACAGGGGGTG	TGGTGACCGA	CTTCGACCGA	GACGGGATGC	TGGACCTCAT	CTTGTCCCCT	960
	GGAGAGTCCA	TGGCTCAGCC	GCTGTCCGTC	TTCCGGGGCA	ATCAGGGCTT	CAACAACAAC	1020
	TGGCTGCGAG	TGGTGCCACG	CACCCGGGTT	GGGGCCTTTG	CCAGGGGAGC	TAAGGTCTGT	1080
85	CTCTACACCA	AGAAGAGTGG	GGCCACCTG	AGGATCATCG	ACGGGGGCTC	AGGCTACCTG	1140
	TGTGAGATGG	AGCCCGTGGC	ACACTTTGGC	CTGGGGAAGG	ATGAAGCCAG	CAGTGTGGAG	1200
	GTGACGTGGC	CAGATGGCAA	GATGGTGAGC	CGGAACGTGG	CCAGCGGGGA	GATGAACCTA	1260



5	GTGCTGGAGA	TCCTCTACCC	CCGGGATGAG	GACACACTTC	AGGACCCAGC	CCCACTGGAG	1320
	ACACCAATGA	ATGCATCCAG	TTCCCATTCG	TGTGCCCTCG	AGACAAGCCC	GTATGTGTCA	1380
	ACACCTATGG	AAGCTACAGG	TGCCGGACCA	ACAAGAAGTG	CAGTCGGGGC	TACGAGCCCA	1440
	ACGAGGATGG	CACAGCCTGC	GTGGGGACTC	TCGGCCAGTC	ACCGGGCCCC	CGCCCCACCA	1500
	CCCCCACC GC	TGCTGCTGCC	ACTGCCGCTG	CTGCTGCCGC	TGCTGGAGCT	GCCACTGCTG	1560
	CACCGGTCTT	CGTAGATGGA	GATCTCAATC	TGGGGTCGGT	GGTTAAGGAG	AGCTGCGAGC	1620
	CCAGCTGCTG	AGCAGGGGTG	GGACATGAAC	CAGCGGATGG	AGTCCAGCAG	GGGAGTGGGA	1680
	AAGTGGGCTT	GTGCTGCTGC	CTAGACAGTA	GGGATGTAAA	GGCCTGGGAG	CTAGACCCTC	1740
10	CCCAAGCCCA	TCCATGCACA	TTACTTAGCT	AACAATTAGG	GAGACTCGTA	AGGCCAGGCC	1800
	CTGTGCTGGG	CACATAGCTG	TGATCACAGC	AGACAGGGTC	GCTGCCCTGA	TGGCGCTTAC	1860
	ATTCCAGTGG	GTCTAATGAC	CATATCTTAG	GACACAGATG	TGCCCAGGGA	GGTGGTGTCA	1920
	CTGCACAGGA	AGTATGAGGA	CTTAGTGTC	CTGAGTTCAA	ATCCTGATTC	AGGAACTCAC	1980
	AAAGCTATGT	GACCTTACAC	CAGTCACTTA	ACTTGTTAGC	CATCCATTAT	CGCATCTGCA	2040
15	AAATGGGGAT	TAAGAATAGA	ATCTTGGGGT	TAGTGTGGAG	ATTAGATTAA	ATGTATGTAA	2100
	GACACTTGGC	ACAAAACCTG	GCACATAGTA	AAGGCTCAAT	AAAAACAAGT	GCCTCTCACT	2160
	GGGCTTTGTC	AACACGTG					

Seq ID NO: 469 Protein sequence  
Protein Accession #: NP\_060528.1

20	1	11	21	31	41	51	
	MDPEASDLRS	GILALRDVAA	EAGVSKYTGG	RGVSVGPILS	SSASDIFCDN	ENGNPFLFHN	60
25	RGDGTFFVDA	ASAGVDDPHQ	HGRGVALADF	NRDGKVDIVY	GNWNGPHRLY	LQMSTHGKVR	120
	FRDIASPKFS	MPSPVRTVIT	ADFDNDQELE	IFFNNIAYRS	SSANRLFRVI	RREHGDPLIE	180
	ELNPGDALEP	EGRGTGGVVT	DFDGDGMLDL	ILSHGESMAQ	PLSVFRGNQG	FNNNWL RVVP	240
	RTRVGAFARG	AKVVLTKKS	GAHLRIIDGG	SGYLCEMEPV	AHFGLGKDEA	SSVEVTWPDG	300
	KMVSRRNVASG	EMNSVLEILY	PRDEDTLQDP	APLETMPNAS	SSHSCALET	PYVSTPMEAT	360
30	GAGPTRSAVG	ATSPTRMAQP	AWGLSASHRA	PAPPPPPLLL	PLPLLLPLLE	LPLLHRSS	

Seq ID NO: 470 DNA sequence  
Nucleic Acid Accession #: AJ279016  
Coding sequence: 1..1962

35	1	11	21	31	41	51	
	ATGTCCAGGA	TGTTACCGTT	CCTGCTGCTG	CTCTGGTTTC	TGCCCCATCAC	TGAGGGGTCC	60
	CAGCGGGCTG	AACCCATGTT	CACTGCAGTC	ACCAACTCAG	TTCTGCCCTCC	TGACTATGAC	120
40	AGTAATCCCA	CCCAGCTCAA	CTATGGTGTG	GCAGTTACTG	ATGTGGACCA	TGATGGGGAC	180
	TTTGAGATCG	TCGTGGCGGG	GTACAATGGA	CCCAACCTGG	TTCTGAAGTA	TGACCGGGCC	240
	CAGAAGCGGC	TGGTGAACAT	CGCGGTCGAT	GAGCGCAGCT	CACCCTACTA	CGCGCTGCGG	300
	GACCGGCAGG	GGAACGCCAT	CGGGGTCACA	GCCTGCGACA	TCGACGGGGA	CGGCCGGGAG	360
	GAGATCTACT	TCCTCAACAC	CAATAATGCC	TTCTCGGGGG	TGGCCACGTA	CACCGACAAG	420
45	TTGTTCAAGT	TCCGCAATAA	CCGGTGGGAA	GACATCCTGA	GCGATGAGGT	CAACGTGGCC	480
	CGTGGTGTGG	CCAGCCTCTT	TGCCGGACGC	TCTGTGGCCT	GTGTGGACAG	AAAGGGCTCT	540
	GGACGCTACT	CTATCTACAT	TGCCAATTAC	GCCTATCTGC	ATGTGGGGCC	TGATGCCCTC	600
	ATTGAAATGG	ACCCTGAGGC	CAGTGACCTC	TCCCGGGGCA	TTCTGGCGCT	CAGAGATGTG	660
	GCTGCTGAGG	CTGGGGTCAG	CAAAATATACA	GGGGGCCGAG	GCGTCAGCGT	GGGCCCCATC	720
50	CTCAGCAGCA	GTGCCTCGGA	TATCTTCTGC	GACAATGAGA	ATGGGCCTAA	CTTCCTTTTC	780
	CACAACCGGG	GCGATGGCAC	CTTGTGTGAC	GCTGCGGCCA	GTGCTGGTGT	GGACGACCCC	840
	CACCAGCATG	GGCGAGGTGT	CGCCCTGGCT	GACTTCAACC	GTGATGGCAA	AGTGGACATC	900
	GTCTATGGCA	ACTGGAATGG	CCCCACCCGC	CTCTATCTGC	AAATGAGCAC	CCATGGGAAG	960
	GTCCGCTTCC	GGGACATCGC	CTCACCAAG	TTCTCCATGC	CCTCCCCTGT	CCGCACGGTC	1020
55	ATCACCGCCG	ACTTTGACAA	TGACCAGGAG	CTGGAGATCT	TCTTCAACAA	CATTGCCTAC	1080
	CGCAGCTCCT	CAGCCAACCG	CCTCTTCCGC	GTCATCCGTA	GAGAGCACGG	AGACCCCTC	1140
	ATCGAGGAGC	TCAATCCCGG	CGACGCCTTG	GAGCCTGAGG	GCCGGGGCAC	AGGGGGTGTG	1200
	GTGACCGACT	TCGACGGAGA	CGGGATGCTG	GACCTCATCT	TGTCCCATGG	AGAGTCCATG	1260
	GCTCAGCCGC	TGTCCGTCTT	CCGGGGCAAT	CAGGGCTTCA	ACAACAACTG	GCTGCGAGTG	1320
60	GTGCCACGCA	CCCGGTTTGG	GGCCTTTGCC	AGGGGAGCTA	AGGTCTGTCT	CTACACCAAG	1380
	AAGAGTGGGG	CCCACCTGAG	GATCATCGAC	GGGGGCTCAG	GCTACCTGTG	TGAGATGGAG	1440
	CCCGTGGCAC	ACTTTGGCCT	GGGGAAGGAT	GAAGCCAGCA	GTGTGGAGGT	GACGTGGCCA	1500
	GATGGCAAGA	TGGTGAGCCG	GAACGTGGCC	AGCGGGGAGA	TGAACCTCAGT	GCTGGAGATC	1560
	CTCTACCCCC	GGGATGAGGA	CACACTTCAG	GACCCAGCCC	CACTGGAGTG	TGGCCAAGGA	1620
65	TTCTCCCAGC	AGGAAAATGG	CCATTGCATG	GACACCAATG	AATGCATCCA	GTTCCCATT	1680
	GTGTGCCCTC	GAGACAAGCC	CTATGTGTGC	AACACCTATG	GAAGCTACAG	GTGCCGGACC	1740
	AACAAGAAGT	GCAGTCGGGG	CTACGAGCCC	AACGAGGATG	GCACAGCCTG	CGTGGGGACT	1800
	CTCGGCCAGT	CACCGGGCCC	CCGCCCCACC	ACCCCCACCG	CTGCTGCTGC	CACTGCCGCT	1860
	GCTGCTGCCG	CTGCTGGAGC	TGCCACTGCT	GCACCGGTCC	TCGTAGATGG	AGATCTCAAT	1920
70	CTGGGGTCGG	TGGTTAAGGA	GAGCTGCGAG	CCCAGCTGCT	GAGCAGGGGT	GGGACATGAA	1980
	CCAGCGGATG	GAGTCCAGCA	GGGGAGTGGG	AAAGTGGGCT	TGTGCTGCTG	CCTAGACAGT	2040
	AGGGATGTAA	AGGCCTGGGA	GCTAGACCC	CCCCAAGCCC	ATCCATGCAC	ATTACTTAGC	2100
	TAACAATTAG	AGGAGCTCGT	AAGGCCAGGC	CCTGTGCTGG	GCACATAGCT	GTGATCACAG	2160
	CAGACAGGGT	CGCTGCCCTG	ATGGCGCTTA	CATTCCAGTG	GGTCTAATGA	CCATATCTTA	2220
75	GGACACAGAT	GTGCCCAGGG	AGGTGGTGTG	ACTGCACAGG	AAGTATGAGG	ACTTTAGTGT	2280
	CCTGAGTTCA	AATCCTGATT	CAGGAACCTCA	CAAAGCTATG	TGACCTTACA	CCAGTCACTT	2340
	AACTTGTTAG	CCATCCATTA	TCGCATCTGC	AAAATGGGGA	TTAAGAATAG	AATCTTGGGG	2400
	TTAGTGTGGA	GATTAGATTA	AATGTATGTA	AGACACTTGG	CACAAAACCT	GGCACATAGT	2460
80	AAAGGCTCAA	TAAAAACAAG	TGCCTCTCAC	TGGCTTTTGT	CAACACG		

Seq ID NO: 471 Protein sequence  
Protein Accession #: CAC08451

85	1	11	21	31	41	51	
	MSRMLPFLLL	LWFLPITEGS	QRAEPMFTAV	TNSVLPPDYD	SNPTQLNYGV	AVTDVDHGDG	60
	FEIVVAGYNG	PNLVLYKIDRA	QKRLVNIADV	ERSSPYALR	DRQGNAGVT	ACDIDGDGRE	120
	EIYFLNTNNA	FSGVATYTDK	LFKFRNNRWE	DILSDEVNVA	RGVASLFAGR	SVACVDRKGS	180



GRYSIYIANY AYGNVGPDAL IEMDPEASDL SRGILALRDV AAEAGVSKYT GGRGVSVGPI 240  
 LSSASDIFC DNENGPNFLF HNRGDGTFVD AAASAGVDDP HQHGRGVALA DFNRDGKVDI 300  
 VYGNWNGPHR LYLQMSHGHK VRFDRDIASPK FSPSPVVRTV ITADFDNDQE LEIFFNNIAY 360  
 RSSSANRLFR VIRREHGDPL IEELNPGDAL EPEGRGTGGV VTDGFDGML DLILSHGESM 420  
 AQPLSVFRGN QGFNNWLRV VPRTRFGAFA RGAKVVLYTK KSGAHLRIID GSGGYLCEME 480  
 PVAHFGLGKD EASSVEVTWP DGKMSVRNVA SGEMNSVLEI LYPRDEDTLQ DPAPLECGQG 540  
 FSQQENGHCHM DTNECIQFPF VCPDRKPVCV NTYGSYRCRT NKKCSRGYEP NEDGTACVGT 600  
 LGQSPGPRPT TPTAAAATAA AAAAGAATA APVLVDGDLN LGSVVKESCE PSC

Seq ID NO: 472 DNA sequence  
 Nucleic Acid Accession #: FGENESHH  
 Coding sequence: 1..4794

1 11 21 31 41 51  
 | | | | | |  
 15 ATGGCGTGTC CGGGAGGACT CCCAGCCCCG TGCTCTGGTT GGATGGGACT GGGTGGGCCC 60  
 AGCGGCTCCT CCCAGCATC CCTCCCCAT TCCTCCTCCA GGTACAATGG ACCCAACCTG 120  
 GTTCTGAAGT ATGACCGGGC CCAGAAGCGG CTGGTGAACA TCGCGGTCGA TGAGCGCAGC 180  
 20 TCACCTACT ACAGCGTGC GACCGGCAG GGAACGCCA TCGGGGTCAC AGCCTGCGAC 240  
 ATCGACGGGG ACAGCGGGGA GGAGATCTAC TTCCTCAACA CCAATAATGC CTTCTCGGGC 300  
 CACAGCAGT CAGCGCAGGT CCTCTGGG CTCCACAGAA ACAGGCCTGT GCTGAAGCCT 360  
 CCACCTACAA CCCCTGCAGG CCTCCTGGGT CTGCCTCCAC TCAGCGGAAG GGACTTTTCC 420  
 TCCTCCCTGG GTCAGGCTTC TCCGGACAGC AGGCAGGGAG AGAGGGTGCC GGTTCCTGTC 480  
 25 TGTCGGGGTG GACTGAGACC TACCATGAA CCAGAACCAT TTCTTCTGAG ACCCAAATCA 540  
 GGGGTGGCCA CGTACACCGA CAAGTTGTTC AAGTTCGCCA ATAACCGGTG GGAAGACATC 600  
 CTGAGCGATG AGGTCAACGT GGCCCGTGGT GTGGCCAGCC TCTTTGCCGG ACCTCTGTG 660  
 GCCTGTGTGG ACAGAAAGGG CTCTGGACGC TACTCTATCT ACATTGCCAA TTACGCCTAC 720  
 GGTAATGTGG GCCCTGATGC CCTCATGAA ATGGACCCTG AGGCCAGTGA CCTCTCCCGG 780  
 30 GGCATTCTGG CGCTCAGAGA TGTGGCTGCT GAGGCTGGGG TCAGCAAATA TACAGAAAGGC 840  
 TTCTCCACA CTGCCTCTCC AAGCATTGGT GAGATATCTG GCAGAACCAG GAGCGGGGAA 900  
 GGAGGAGACC CAGAGGAGGC AGATGAGGAG CACAGTGGGG ATGGAAGCAC CAGCCAACTG 960  
 TGCCGGCTGG GCTGGAAGGA CGGGCAGTTC AAGGAAGAAG CAGCAGCTTT GGTGGAGGAA 1020  
 CAGAGGGAGG CTGGGGCAGC TGGCGTGCCC AGAGACGCTG TTCGAACAGC TCTGCAGACT 1080  
 35 TCCAAAAGCC ATTTGGCTGA CAAGAACCTA TTTGGCCAC CATGTTACTA TTCTGTCTGC 1140  
 GCGCCTTCTC CAGCCCACCC TTTCCCTGCC CGCCAAGCCC CCCAACACTA CCCTGTAGCC 1200  
 CCCCTTGTC CTAGCTAAT GACACATGGA CGTCTGGCTG GAAAACCTAGC CCGGAGTGTC 1260  
 CCCCACCCCC GAGCCCCAGG AATGGACCCC AAATGTAAGG GCCGCCATGC TGAGCCCCGGC 1320  
 CTGATGGCTG AGGCTTTGGG CGCGTGCCCA GCGCTCAGCA CCACTGTGGT GCCAGGGGGC 1380  
 40 CTGAGAAGCT GGGAGGAAAG CAGGCAGAA GGGCAGGCCA TGTCCAGATG TGCACTCAGG 1440  
 GAGCTGGGAG GTCCCTGAG CCAAGCTACA CAGCACTGC CTGCTAGAGA GCTGTATGAC 1500  
 CTGGGAGAAC CTCCATTTT ACAAGAACA GACGGAGATC CAGGGAGGAG AAGGGACTCG 1560  
 CCCAAGGTCA CACAGGAGTG CCATCTAGTG GCCACCATGC CAGCTCTCGG GGGACTCGAG 1620  
 GGCCCCGGGA GGGTGGCCAA GCGAGAGATT GGGAGAGAGA CTGGGGCAGT AGGAAGACCA 1680  
 45 CTCTCCCATC CCCTGGTCCC CAACCTCCCC AGCTGCTTGA GGCTCTTGA AGCCGGGACA 1740  
 GTGCCGGGAG CTGCCCTGCC TGGGAATCCT GGGAACTGGG TTCTGGACAT GGCCAAGGCC 1800  
 CTGGCGTGGA ACCAGATGGA AAAAGAGGAG GGAAGATTC ATGGAGACCA TGAGCCCAGA 1860  
 TTTAGGCTCA GGAAGCACG GGAAGCAGAA TTCCCCCAG GCTCCTCTGA GGAGCCTCTG 1920  
 CTGCAGTTCC CCTCAGGCCT CAGAGGCAGC CCTGTCTCC AGGTGGGCCT GGGGCTTGCT 1980  
 50 TCTGCCACTC ACTGTGGGTC GATGTCTTTT CTAGGGGGCC GAGGCGTCAG CGTGGGCCCC 2040  
 ATCCTCAGCA GCAGTGCCCT GGATATCTTC TGCGACAATG AGAATGGGCC TAACTTCCTT 2100  
 TTCCACAACC GGGGCGATGG CACCTTTGTG GACGCTGCGG CCAGTGCTGA ACGTCGTTTA 2160  
 GCCTTCATCG TTCACCTCAA ATATCACCTC TGCAGAGATT TTCTCACTC CTTGTGCCAC 2220  
 CTAGCAGAAA CTGGTCCCTC CTCCTCTGTC TGCCCGTGGC ATGCACGTCT TCTTCAGGCT 2280  
 55 CCACATTGCC ATCATGGTTT GTCTATGAGC TTTACAAGGA CCGGGTCACG GTTCTATTCA 2340  
 TTCTTGACGC AAGGCTTGGC CTCCAGTGCC CACCGGAGGA CACTCAGCCT CCAGGGTTCT 2400  
 CAGGGGGCCC CACCCTGCCT TCTGGCAAGA GCTCCCTGTG TCCTGGGGTC TCTGATCCCC 2460  
 ACTGCCTATT ACATTGTCTT GTGGTCTGCC ATCCCAGAGA GCCTGATGAC CCACAGCTAT 2520  
 TTGTCTCTG AAAGAGTCAA CGTGGGTGTG GACGACCCCC ACCAGCATGG GCGAGGTGTC 2580  
 60 GCCCTGGCTG ACTTCAACCG TGATGGCAAA GTGCACATCG TCTATGGCAA CTGGAATGGC 2640  
 CCCCACCGCC TCTATCTGCA AATGAGCACC CATGGGAAGG TCCGCTTCCG GGACATCGCC 2700  
 TCACCCAAGT TCTCCATGCC CTCCCCTGTC CGCACGGTCA TCACCGCCGA CTTTGACAAT 2760  
 GACCAGGAGC TGGAGATCTT CTTCAACAAC ATTGCCTACC GCAGCTCCTC AGCCAACCGC 2820  
 CTCTTCCGAT GCTCCATCCT GGCTCGTGGC TCTTCATCCT TGACAGCTGG TGGGAGGAAC 2880  
 65 GGTCAAGGAG AAGGTTTAAG AATCAGAAGG GGAAGGTTCC CAGGGCCAGG GGGTCAGGCC 2940  
 AAGGTCAACA CAGGTCCCTT GATGAAGAAA CAGAAGGAA GGAAGGACGA GGACTGGGCA 3000  
 AGAGGCTGTG GGAATGCAGG GCAAAGCCTG GCCAAGGAGC CGGCTCTGCT TATTGCAGGG 3060  
 AAAGGGAAGG GAAATGTGGC CCAAAGTGTG CCCAGAACC AAGCGCCACA AGATACAAAG 3120  
 CCACACTACC ACAAAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAAG GGGCTACGGG 3180  
 70 GTCCAATCAC TACCAGGAAA AGGGGCTACG GGGTCCAATC ACTACCAGGA AAAGGGGCTA 3240  
 CGGGGTCCAA TCACTACCAG GAAAAGGGGC TACGGGGTCC AATCACTACC AGGAAAAGGG 3300  
 GCTACGGGCT CCAATCACTA CCAGGAAAAG GGGCTACAGG GTCCAATCAC TACCAGGAAA 3360  
 AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA CAGGGTCCAA TCACTACCAC 3420  
 AGAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGGG GCTACGGGGT CCAATCACTA 3480  
 75 CCAGGAAAAG GGGCTACAGG GTCCAATCAC TACCAGGAAA AGGGGCTACG GGGTCCAATC 3540  
 ACTACCAGGA AAAGGGGCTA CGGGCTCCAA TCACTACCAG GAAAAGGGGC TACGGGGTCC 3600  
 AATCACTACC AGGAAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAAG GGGCTACAGG 3660  
 GTCCAATCAC TACCACAGAA AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA 3720  
 CGGGGTCCAA TCACTACCAG GAAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGAG 3780  
 80 GCTATGGGGT CCAATCACTA CCAGGAAAAG GGGCTACGGG CTCCAATCAC TACCAGGAAA 3840  
 AGGGGCTATG GGGTCCAATC ACTACCACAG AAAGGGGCTA CGGGGTCCAA CGTCATCCGT 3900  
 AGAGAGCACG GAGACCCCTT CATCGAGGAG CTCAATCCCG GCGACGCCCT GGAGCCTGAG 3960  
 GGCCGGGGCA CAGGGGGTGT GTGACCGGAC TTCGACGGAG ACGGGATGCT GGACCTCATC 4020  
 TTGTCCCATG GAGAGTCCAT GGCTCAGCCG CTGTCGGTCT TCCGGGGCAA TCAGGGCTTC 4080  
 85 AACAACTAAGT GGCTGCGAGT GGTGCGCAGC ACCGCTTTTG GGGCTTTTGC CAGGGGAGCT 4140  
 AAGGTCGTGC TCTACACCAA GAAGAGTGGG GCCACCTGA GGATCATCGA CGGGGGCTCA 4200  
 GGCTACCTGT GTGAGATGGA GCGGCTGGCA CACTTTGGCC TGGGGAAGGA TGAAGCCAGC 4260  
 AGTGTGGAGG TGACGTGGCC AGATGGCAAG ATGGTGAGCC GGAACGTGGC CAGCGGGGAG 4320



WO 02/086443

PCT/US02/12476

ATGAACCTCAG TGCTGGAGAT CCTCTACCCC CGGGATGAGG ACACACTTCA GGACCCAGCC 4380  
 CCACTGGAGT GTGGCCAAGG ATTCTCCCAG CAGGAAAATG GCCATTGCAT GGACACCAAT 4440  
 GAATGCATCC AGTTCCCAT TCGTGTGCCCT CGAGACAAGC CCGTATGTGT CAACACCTAT 4500  
 GGAAGCTACA GGTGCCGGAC CAACAAGAAG TGCAGTCGGG GCTACGAGCC CAACGAGGAT 4560  
 GGCACAGCCT GCGTGGGTAC TGAGCTAGGC TCTAGGCATA CAATGACGTG GAAACCAAGG 4620  
 CCCAAAAGG AGCTGCAACT TTCCCAAGGC ATCTGCACCC CCGTCTGGTC CTTTTTCTTG 4680  
 CCGGGTTGCC GGCTGCTCCT CAAAAGAGCT CAGCTCCAGG CTGCTCCAG CACCTTCTC 4740  
 CAGAAAGCTC CAGGTATTCC AGAAGCCCCA GTGTATGAAC AAGATCAGGA ATAA

Seq ID NO: 473 Protein sequence  
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51  
 | | | | |  
 MACPGGLPAR CSGWMGLGGP SGSSPASPPH SSSRYNGPNL VLKYDRAQKR LVNIAVDERS 60  
 SPYYALRDRQ GNAIGVTACD IDGDGREETIY FLNTNNAFSG HSSSAQVPSG LHRNRPVLKP 120  
 PPTTPAGLLG LPPLSGRDFS SSLGQASPDG RQGERVPVPC CRGGLRPTHE PEPFLLRPKS 180  
 GVATYTDKLF KFRNNRWEDI LSDEVNVARG VASLFAGRSV ACVDRKGSGR YSIYIANYAY 240  
 GNVGPDALIE MDPEASDLR GILALRDVAA EAGVSKYTEG FSHTASPSIG EISGRTEERE 300  
 GGDPEEADDEE HSGDGTSQL CRLGWKDGQF KEEAALVEE QREAGAAGVP RGRVRTALQT 360  
 SKSHLADKNL FGPPCYYSVC APSAPHPFPA RQAPQHYVPA PLVTQLMTHG RLAGKLARSV 420  
 PHPRAPGMDP KCKGRHAEPG LMAEALGAWP ALSTTVVPGG LRSWEESRQK GQAMSRCLAR 480  
 ELGGPWSQAT QHLPARELYD LGEPPILOQT DGDGPRRRDS PKVTQECHLV ATMPALGGLE 540  
 GPGRVAKREI GRETGAVGRP LSHPLVPNFP SCLRPLEAGT VPGAALPGNP GNWVLDMAKA 600  
 LAWNQMEKEE GKIHGDEPR FRLRKAREAE FPPGSSEEP LQFPSGLRGS PVLQVGLGLA 660  
 SATHCGSMSF LGGRGVSVGP ILSSASDIF CDNENGNPFL FHNRGDGTFFV DAAASAERRL 720  
 AFIVHLKYHL CRDFPHSLCH LAETGPSSSC CPWHARLLQA PHCHHGLSMS FTRTGSRFYS 780  
 FLTQGLASSA HRRTLSLQGS QGAPPCLLAR APCVLGSLIP TAYYIVLWSA IPESLMTHSY 840  
 LSSERVNVGV DDPHQHGRGV ALADFNDRDG VDIVYGNWNG PHRLYLQMSH HGKVRFRDIA 900  
 SPKFSMPSPV RTVITADFDN DQELEIFFNN IAYRSSSANR LFRCSILARG SSSLTAGGRN 960  
 GQGEGLRIRR GGFPGPGGQA KVNTPGLMKK QKGRKDEDAW RGCGNAGQSL AKEPASAIAG 1020  
 KGKGNVAQSV PRTQAPQDTK PHYHKKGLQG PITTRKRGYV VQSLPGKGAT GSNHYQEKGL 1080  
 RGPITTRKRG YGVQSLPGKG ATGSNHYQEK GLQGPITTRK RGYGLQSLPG KGATGSNHYH 1140  
 RKGLRAPITT RKRGVGVQSL PGKGATGSNH YQEKGLRGPI TTRKRGYGLQ SLPGKGATGS 1200  
 NHYQEKGLQG PITTRKRGYR VQSLPGKGAT GSNHYQEKGL RGPITTRKRG YGLQSLPGKE 1260  
 AMGSNHYQEK GLRAPITTRK RGYGVQSLPG KGATGSNVIR REHGDPLIEE LNPGDALPEE 1320  
 GRGTGGVVD FDGDGMLDLI LSHGESMAQP LSVFRGNQGF NNNWLRVVR TRFGAFARGA 1380  
 KVVLYTKKSG AHLRIIDGGS GYLCEMEPVA HFGLKDEAS SVEVTWPDGK MVSRRNVASGE 1440  
 MNSVLEILYP RDEDTLQDPA PLECCQGFSP QENGHCMETN ECIQFPFVCP RDKPVCVNTY 1500  
 GSYRCRTNKK CSRGYEPNED GTACVGTGLG SRHTMTWKPR PKKELQLSQG ICTPVWSFFL 1560  
 PGCRLLLKRA QLQAAPSTLL QKAPGIPEAQ VYEQDQE

Seq ID NO: 474 DNA sequence  
 Nucleic Acid Accession #: NM\_003661.1  
 Coding sequence: 1..1152

1 11 21 31 41 51  
 | | | | |  
 ATGAGTGCAC TTTTCCTTGG TGTGGGAGTG AGGGCAGAGG AAGCTGGAGC GAGGGTGCAA 60  
 CAAACGTTTC CAAGTGGGAC AGATACTGGA GATCCTCAAA GTAAGCCCTT CCGTGACTGG 120  
 GCTGCTGGCA CCATGGACCC AGAGAGCAGT ATCTTTATTG AGGATGCCAT TAAGTATTTT 180  
 AAGGAAAAAG TGAGCACACA GAATCTGCTA CTCCTGCTGA CTGATAATGA GGCCTGGAAC 240  
 GGATTCGTGG CTGCTGCTGA ACTGCCAGG AATGAGGCAG ATGAGCTCCG TAAAGCTCTG 300  
 GACAACCTTG CAAGACAAAT GATCATGAAA GACAAAAACT GGCACGATAA AGGCCAGCAG 360  
 TACAGAAACT GGTTCCTGAA AGAGTTTCTT CCGTTGAAAA GTGAGCTTGA GGATAACATA 420  
 AGAAGGCTCC GTGCCCTTGC AGATGGGGTT CAGAAGGTCC ACAAAGGCAC CACCATCGCC 480  
 AATGTGGTGT CTGGCTCTCT CAGCATTTCC TCTGGCATCC TGACCCTCGT CCGCATGGGT 540  
 CTGGCACCTT TCACAGAGGG AGGCAGCCTT GTACTCTTGG AACCTGGGAT GGAGTTGGGA 600  
 ATCACAGCCG CTTTGACCGG GATTACCAGC AGTACCATGG ACTACGGAAG GAAGTGGTGG 660  
 ACACAAGCCC AAGCCCACGA CCTGGTCAAT AAAAGCCTTG ACAAATTGAA GGAGGTGAGG 720  
 GAGTTTTTGG GTGAGAACAT ATCCAACCTT CTTTCCTTAG CTGGCAATAC TTACCAACTC 780  
 ACACGAGGCA TTGGGAAGGA CATCCGTGCC CTCAGACGAG CCAGAGCCAA TCTTCAGTCA 840  
 GTACCGCATG CCTCAGCCTC ACGCCCCCGG GTCACAGAG CAATCTCAGC TGAAGCGGT 900  
 GAACAGGTGG AGAGGGTTAA TGAACCCAG ATCCTGGAAA TGAGCAGAGG AGTCAAGCTC 960  
 ACGGATGTGG CCCCTGTAAG CTTCTTTCTT GTGCTGGATG TAGTCTACCT CGTGTACGAA 1020  
 TCAAAGCACT TACATGAGGG GGCAAAGTCA GAGACAGCTG AGGAGCTGAA GAAGGTGGCT 1080  
 CAGGAGCTGG AGGAGAAGCT AAACATTCTC AACAATAATT ATAAGATTCT GCAGGCGGAC 1140  
 CAAGAAGTGT GA

Seq ID NO: 475 Protein sequence  
 Protein Accession #: NP\_003652.1

1 11 21 31 41 51  
 | | | | |  
 MSALFLGVGV RAEEAGARVQ QNVPSGTDGT DPQSKPLGDW AAGTMDPES IFIEDAIKYF 60  
 KEKVSTQNL LLLTDNEAWN GFVAAAELEP NEADELRKAL DNLRQMIMK DKNWHDKGQQ 120  
 YRNWFLKEFP RLKSELEDNI RRLRALADGV QKVHKGTTIA NVVSGSLSS SGILTLVGMG 180  
 LAPFTEGGSL VLLPEGMELG ITAALTGITS STMDYGKKWW TQAQAHDLVI KSLDKLKEVR 240  
 EFLGENISNF LSLAGNTYQL TRGIGKDIRA LRRARANLQS VPHASASRPR VTEPISAESG 300  
 EQVERVNEPS ILEMSRGVKL TDVAPVSFFL VLDVVYLVYE SKHLHEGAKS ETAEELKKVA 360  
 QELEEKLNL NNNYKILQAD QEL

Seq ID NO: 476 DNA sequence  
 Nucleic Acid Accession #: NM\_014452.1  
 Coding sequence: 1..1968

1 11 21 31 41 51



	ATGGGGACCT	CTCCGAGCAG	CAGCACCGCC	CTCGCCTCCT	GCAGCCGCAT	CGCCCGCCGA	60
	GCCACAGCCA	CGATGATCGC	GGGCTCCCTT	CTCCTGCTTG	GATTCCCTAG	CACCACCACA	120
5	GCTCAGCCAG	AACAGAAGGC	CTCGAATCTC	ATTGGCACAT	ACCGCCATGT	TGACCGTGCC	180
	ACCGGCCAGG	TGCTAACCTG	TGACAACTGT	CCAGCAGGAA	CCTATGTCTC	TGAGCATTGT	240
	ACCAACACAA	GCCTGCGCGT	CTGCAGCAGT	TGCCCTGTGG	GGACCTTTAC	CAGGCATGAG	300
	AATGGCATAG	AGAAATGCCA	TGACTGTAGT	CAGCCATGCC	CATGGCCAAT	GATTGAGAAA	360
	TTACCTTTGTG	CTGCCCTTAC	TGACCGAGAA	TGCACTTGCC	CACCTGGCAT	GTTCCAGTCT	420
10	AACGCTACCT	GTGCCCCCCA	TACGGTGTGT	CCTGTGGGTT	GGGGTGTGCG	GAAGAAAGGG	480
	ACAGAGACTG	AGGATGTGCG	GTGTAAGCAG	TGTGCTCGGG	GTACCTTCTC	AGATGTGCCT	540
	TCTAGTGTGA	TGAAATGCAA	AGCATACACA	GACTGTCTGA	GTCAGAACCT	GGTGGTGATC	600
	AAGCCGGGGA	CCAAGGAGAC	AGACAACGTC	TGTGGCACAC	TCCCGTCCCT	CTCCAGCTCC	660
	ACCTCACCTT	CCCCTGGCAC	AGCCATCTTT	CCACGCCCTG	AGCACATGGA	AACCCATGAA	720
15	GTCCCTTCCT	CCACTTATGT	TCCCAAAGGC	ATGAACTCAA	CAGAATCCAA	CTCTTCTGCC	780
	TCTGTTAGAC	CAAAGGTACT	GAGTAGCATC	CAGGAAGGGA	CAGTCCCTGA	CAACACAAGC	840
	TCAGCAAGGG	GGAAGGAAGA	CGTGAACAAG	ACCTTCCCAA	ACCTTCAGGT	AGTCAACCAC	900
	CAGCAAGGCC	CCCACCACAG	ACACATCCTG	AAGCTGCTGC	CGTCCATGGA	GGCCACTGGG	960
	GGCGAGAAGT	CCAGCACGCC	CATCAAGGGC	CCCAAGAGGG	GACATCCTAG	ACAGAACCTA	1020
20	CACAAGCATT	TTGACATCAA	TGAGCATTTC	CCCTGGATGA	TTGTGCTTTT	CCTGCTGCTG	1080
	GTGCTTGTGG	TGATTGTGGT	GTGCAGTATC	CGGAAAAGCT	CGAGGACTCT	GAAAAAGGGG	1140
	CCCCGGCAGG	ATCCCACTGC	CATTGTGGAA	AAGGCAGGGC	TGAAGAAATC	CATGACTCCA	1200
	ACCCAGAACC	GGGAGAAATG	GATCTACTAC	TGCAATGGCC	ATGGTATCGA	TATCCTGAAG	1260
	CTTGTTAGCAG	CCCAAGTGGG	AAGCCAGTGG	AAAGATATCT	ATCAGTTTCT	TTGCAATGCC	1320
25	AGTGAGAGGG	AGGTTGCTGC	TTTCTCCAAT	GGGTACACAG	CCGACCACGA	GCGGGCCTAC	1380
	GCAGCTCTGC	AGCACTGGAC	CATCCGGGGC	CCCGAGGCCA	GCCTCGCCCA	GCTAATTAGC	1440
	GCCCTGCGCC	AGCACCGGAG	AAACGATGTT	GTGGAGAAGA	TTCTGTGGGT	GATGGAAGAC	1500
	ACCACCCAGC	TGGAAACTGA	CAAACCTAGT	CTCCCGATGA	GCCCCAGCCC	GCTTAGCCCC	1560
	AGCCCCATCC	CCAGCCCCAA	CGCGAAACTT	GAGAATTCCG	CTCTCCTGAC	GGTGGAGCCT	1620
30	TCCCCACAGG	ACAAGAACAA	GGGCTTCTTC	GTGGATGAGT	CGGAGCCCCT	TCTCCGCTGT	1680
	GACTCTACAT	CCAGCGGCTC	CTCCCGCTGC	AGCAGGAACG	GTTCCTTTAT	TACCAAAGAA	1740
	AAGAAGGACA	CAGTGTTCGC	GCAGGTACGC	CTGGACCCCT	GTGACTTGCA	GCCTATCTTT	1800
	GATGACATGC	TCCACTTTCT	AAATCCTGAG	GAGCTGCGGG	TGATTGAAGA	GATCCCCCAG	1860
	GCTGAGGACA	AACTAGACCG	GCTATTGCAA	ATTATTGGAG	TCAAGAGCCA	GGAAGCCAGC	1920
35	CAGACCCTCC	TGGACTCTGT	TTATAGCCAT	CTTCCTGACC	TGCTGTAG		

Seq ID NO: 477 Protein sequence  
Protein Accession #: NP\_055267.1

40	1	11	21	31	41	51	
	MGTSPPSSSTA	LASCSRIARR	ATATMIAGSL	LLLGLFLSTTT	AQPEQKASNL	IGTYRHVDRA	60
	TGQVLTCDC	PAGTYVSEHC	TNTSLRVCSS	CPVGTFTTRHE	NGIEKCHDCS	QPCPWPMIEK	120
	LPCAALTDRE	CTCPPGMFQS	NATCAPHTVC	PVGWGVRRKK	TETEDVRCKQ	CARGTFSDVP	180
45	SSVMKCKAYT	DCLSQNLVVI	KPGTKETDNV	CGTLPSFSSS	TSPSPGTAIF	PRPEHMETHE	240
	VPSTYVPK	MNSTESNSSA	SVRPKVLSSI	QEGTVPDNTS	SARGKEDVNK	TLPLNLQVNH	300
	QQGPHRRHIL	KLLPSMEATG	GEKSSTPIKG	PKRGHPRQNL	HKHFDINEHL	PWMIVLFLLL	360
	VLVVIVVCSI	RKSSRTLKKG	PRQDPSAIVE	KAGLKKSMTP	TQNREKWIYY	CNGHGIDILK	420
	LVAAQVGSQW	KDIYQFLCNA	SEREVAAFSN	GYTADHERAY	AALQHWTRIG	PEASLAQLIS	480
50	ALRQHRRNDV	VEKIRGLMED	TTQLETDKLA	LPMSPSPPLSP	SPIPSNNAKL	ENSALLTVEP	540
	SPQDKNKGFF	VDESEPLLR	DSTSSGSSAL	SRNGSFITKE	KKDITVLRQVR	LDPDCLQPIF	600
	DDMLHFLNPE	ELRVIEEIPQ	AEDKLDRLFE	IIGVKSQEAS	QTLSDSVYSH	LPDLL	

Seq ID NO: 478 DNA sequence  
Nucleic Acid Accession #: XM\_044533  
Coding sequence: 238..2751

60	1	11	21	31	41	51	
	GCTCTGCCCA	AGCCGAGGCT	GCGGGGCCGG	CGCCGGCCGG	AGGACTGCGG	TGCCCCGCGG	60
	AGGGGCTGAG	TTTGCCAGGG	CCCACTTGAC	CCTGTTTCCC	ACCTCCCGCC	CCCCAGGTCC	120
	GGAGGCGGGG	GCCCCCGGGG	CGACTCGGGG	GCGGACCGCG	GGGCGGAGCT	GCCGCCCCGTG	180
	AGTCCGGCCG	AGCCACCTGA	GCCCCGAGCC	CGGGACACCG	TCGCTCCTGC	TCTCCGAATG	240
	CTGCGCACCG	CGATGGGCC	GAGGAGCTGG	CTCGCCGCCC	CATGGGGCGC	GCTGCCGCCT	300
65	CGGCCACCGC	TGCTGTGCT	CCTGCTGCTG	TGCTGCTGCT	TGCAAGCCGC	GCCTCCGACC	360
	TGGGCGCTCA	GCCCCCGGAT	CAGCCTGCCT	CTGGGCTCTG	AAGAGCGGCC	ATTCTTCAGA	420
	TTCAAGCTG	AACACATCTC	CAACTACACA	GCCCTTCTGC	TGAGCAGGGA	TGGCAGGACC	480
	CTGTACGTGG	GTGCTCGAGA	GGCCCTCTTT	GCACTCAGTA	GCAACCTCAG	CTTCTTGCCA	540
	GGCGGGGAGT	ACCAGGAGCT	GCTTTGGGGT	GCAGACGCAG	AGAAGAAACA	GCAGTGCAGC	600
70	TTCAAGGGCA	AGGACCCACA	GCGCGACTGT	CAAACTACAT	TCAAGATCCT	CCTGCCGCTC	660
	AGCGGCAGTC	ACCTGTTTAC	CTGTGGCACA	GCAGCCTTCA	GCCCCATGTG	TACCTACATC	720
	AACATGGAGA	ACTTCACCCT	GGCAAGGGAC	GAGAAGGGGA	ATGTCTCTCT	GGAAGATGGC	780
	AAGGGCCGTT	GTCCCTTCGA	CCCGAATTTT	AAGTCCACTG	CCCTGGTGGT	TGATGGCGAG	840
	CTCTACACTG	GAACAGTCAG	CAGCTTCCAA	GGGAATGACC	CGGCCATCTC	GCGGAGCCAA	900
75	AGCCTTCGCC	CCACCAAGAC	CGAGAGCTCC	CTCAACTGGC	TGCAAGACCC	AGCTTTTGTG	960
	GCCTCAGCCT	ACATTCTCTG	GAGCCTGGGG	AGCTTGCAAG	GCGATGATGA	CAAGATCTAC	1020
	TTTTTCTTCA	GCGAGACTGG	CCAGGAATTT	GAGTTCTTTG	AGAACACCAT	TGTGTCCCGC	1080
	ATTGCCCCGA	TCTGCAAGGG	CGATGAGGGT	GGAGAGCGGG	TGCTACAGCA	GCGCTGGACC	1140
	TCCTTCTCTA	AGGCCAGCT	GCTGTGCTCA	CGGCCCGACG	ATGGCTTCCC	CTTCAACGTG	1200
80	CTGCAGGATG	TCTTCACGCT	GAGCCCCAGC	CCCCAGGACT	GGCGTGACAC	CCTTTTCTAT	1260
	GGGGTCTTCA	CTTCCCAGTG	GCACAGGGGA	ACTACAGAAG	GCTCTGCCGT	CTGTGTCTTC	1320
	ACAATGAAGG	ATGTGCAGAG	AGTCTTCAGC	GGCCTCTACA	AGGAGGTGAA	CCGTGAGACA	1380
	CAGCAGTGCT	ACACCGTGAC	CCACCCGGTG	CCCACACCCC	GGCCTGGAGC	GTGCATCACC	1440
	AACAGTGCCC	GGGAAAGGAA	GATCAACTCA	TCCCTGACGC	TCCAGACCCG	CGTGCTGAAC	1500
85	TTCTCAAGG	ACCACTTCCT	GATGGACGGG	CAGGTCCGAA	GCCGCATGCT	GCTGCTGCAG	1560
	CCCCAGGCTC	GCTACCAGCG	CGTGGCTGTA	CACCGCGTCC	CTGGCCTGCA	CCACACCTAC	1620
	GATGTCTCT	TCCTGGGCAC	TGGTGACGGC	CGGCTCCACA	AGGCAGTGAG	CGTGGGCCCC	1680
	CGGGTGCACA	TCATTGAGGA	GCTGCAGATC	TTCTCATCGG	GACAGCCCGT	GCAGAATCTG	1740



5  
10  
15  
20  
25  
30  
35

CTCCTGGACA CCCACAGGGG GCTGCTGTAT GCGGCCTCAC ACTCGGGCGT AGTCCAGGTG 1800  
CCCATGGCCA ACTGCAGCCT GTACAGGAGC TGTGGGGACT GCCTCCTCGC CCGGGACCCC 1860  
TACTGTGCTT GGAGCGGCTC CAGCTGCAAG CACGTGAGCC TCTACCAGCC TCAGCTGGCC 1920  
ACCAGGCCGT GGATCCAGGA CATCGAGGGA GCCAGCGCCA AGGACCTTTG CAGCGCGTCT 1980  
TCGGTTGTGT CCCCCTCTTT TGTACCAACA GGGGAGAAGC CATGTGAGCA AGTCCAGTTC 2040  
CAGCCCAACA CAGTGAACAC TTTGGCCTGC CCGCTCCTCT CCAACCTGGC GACCCGACTC 2100  
TGGCTACGCA ACGGGGCCCC CGTCAATGCC TCGGCCTCCT GCCACGTGCT ACCCACTGGG 2160  
GACCTGCTGC TGGTGGGCAC CCAACAGCTG GGGGAGTTCC AGTGCTGGTC ACTAGAGGAG 2220  
GGCTTCCAGC AGCTGGTAGC CAGCTAGTGC CCAGAGGTGG TGGAGGACGG GGTGGCAGAC 2280  
CAAACAGATG AGGGTGGCAG TGTACCCGTC ATTATCAGCA CATCGCGTGT GAGTGCACCA 2340  
GCTGGTGGCA AGGCCAGCTG GGGTGCAGAC AGGTCTTACT GGAAGGAGTT CCTGGTGATG 2400  
TGCACGCTCT TTTGTGCTGGC CGTGCTGCTC CCAGTTTTAT TCTTGCTCTA CCGGCACCGG 2460  
AACAGCATGA AAGTCTTCCT GAAGCAGGGG GAATGTGCCA GCGTGACCCC CAAGACCTGC 2520  
CCTGTGGTGC TGCCCCCTGA GACCCGCCCC CTCAACGGCC TAGGGCCCCC TAGCACCCCG 2580  
CTCGATCACC GAGGGTACCA GTCCCTGTCA GACAGCCCCC CGGGGTCCCG AGTCTTCACT 2640  
GAGTCAGAGA AGAGGCCACT CAGCATCCAA GACAGCTTCG TGGAGGTATC CCCAGTGTGC 2700  
CCCCGGCCCC GGGTCCGCCT TGGCTCGGAG ATCCGTGACT CTGTGGTGTG AGAGCTGACT 2760  
TCCAGAGGAC GCTGCCCTGG CTTAGGGGCG TGTGAATGCT CGGAGAGGGT CAACTGGACC 2820  
TCCCCTCCGC TCTGCTCTTC GTGGAACACG ACCGTGGTGC CCGGCCCTTG GGAGCCTTGG 2880  
GGCCAGCTGG CCTGCTGCTC TCCAGTCAAG TAGCGAAGCT CCTACCACCC AGACACCCAA 2940  
ACAGCCGTGG CCCCAGAGGT CCTGGCCAAA TATGGGGGCC TGCCTAGGTT GGTGGAACAG 3000  
TGCTCCTTAT GTAAACTGAG CCTTTTGTAT AAAAACAAT TCCAAATGTG AAATAAGAAT 3060  
GAGAGGGAAG AGATAGCATG GCATGCAGCA CACACGGCTG CTCCAGTTCA TGGCCTCCCA 3120  
GGGGTGTGCG GGATGCATCC AAAGTGGTTG TCTGAGACAG AGTTGGAAC CCTCACCAAC 3180  
TGGCCTCTTC ACCTTCCACA TTATCCCGCT GCCACCGGCT GCCCTGTCTC ACTGCAGATT 3240  
CAGGACCAGC TTGGGCTGCG TCGCTTCTGC CTTGCCAGTC AGCCGAGGAT GTAGTTGTTG 3300  
CTGCCGTCGT CCCACCACCT CAGGGACCAG AGGGCTAGGT TGGCACTGCG GCCCTACCA 3360  
GGTCTTGGGC TCGGACCCAA CTCTGGACC TTTCCAGCCT GTATCAGGCT GTGCGCACAC 3420  
GAGAGACAG CGCGAGCTCA GGAGAGATTT CGTGACAATG TACGCCCTTC CCTCAGAATT 3480  
CAGGGAAGAG ACTGTCGCCT GCCTTCTCTC GTTGTGCGT GAGAACCCGT GTGCCCTTC 3540  
CCACCATATC CACCCTCGCT CCATCTTTGA ACTCAAACAC GAGGAACATA CTGCACCTG 3600  
GTCTCTCTCC CAGTCCCCAG TTCACTCTCC ATCCCTCACC TTCTTCCACT CTAAGGGATA 3660  
TCAACACTGC CCAGCACAGG GGCCCTGAAT TTATGTGGTT TTTATACATT TTTTAATAAG 3720  
ATGCACCTTA TGTCACTTTT TAATAAAGTC TGAAGAATTA CTGTTT

Seq ID NO: 479 Protein sequence  
Protein Accession #: XP\_044533.3

40  
45  
50  
55

1 11 21 31 41 51  
| | | | |  
MLRTAMGLRS WLAAPWGALP PRPPLLLLLL LLLLLQPPPP TWALSPRISL PLGSEERPFL 60  
RFEAEHISNY TALLSRDGR TLYVGAREAL FALSSNLSFL PGGEYQELLW GADAEKKQQC 120  
SFKGKDPQRD CQNYIKILLP LSGSHLFTCG TAAFSMPCTY INMENFTLAR DEKGNVLLED 180  
GKGRCPFDPN FKSTALVVDG ELYTGTVSSF QGNDPAISRS QSLRPTKTES SLNWLQDPAF 240  
VASAYIPESL GSLQGDDEKI YFFSETGQE FEFFENTIVS RIARICKGDE GGERVLQQRW 300  
TSFLKAQLLC SRPDDGFPPN VLQDVFTLSP SPQDWRDTLF YGVFTSQWHR GTEGSAVCV 360  
FTMKDVQRFV SGLYKEVNRE TQQWYTVTHP VTPRPGACI TNSARERKIN SSLQLPDRVL 420  
NFLKDHFLMD GQVRSRMLLL QPQARYQOVA VHRVPGLHHT YDVLFLGTGD GRLHKAIVSVG 480  
PRVHIIIEELQ IFSSGQPVQN LLLDTHRGLL YAASHSGVVQ VPMANCSLYR SCGDCLLARD 540  
PYCAWSGSSC KHVSLYQPQL ATRPWIQDIE GASAKDLCSA SSVVSPSFVP TGEKPCQVQ 600  
FQPNVTNTLA CPLLSNLATR LWLRNGAPVN ASASCHVLPT GDLLLVTGQQ LGEFQCWSLE 660  
EGFQQLVASV CPEVVEDGVA DQDEGGSVP VIISTSRVSA PAGGKASWGA DRSYWKFLV 720  
MCTLFVLAVL LPVLFLLYRH RNSMKVFLKQ GECASVHPKT CPVVLPPETR PLNGLGPPST 780  
PLDHRGYQSL SDSPPGSRVF TESEKRPLSI QDSFVEVSPV CPRPRVRLGS EIRDSVV

Seq ID NO: 480 DNA sequence  
Nucleic Acid Accession #: NM\_004217.1  
Coding sequence: 58..1092

60  
65  
70  
75  
80

1 11 21 31 41 51  
| | | | |  
GGCCGGGAGA GTAGCAGTGC CTTGGACCCC AGCTCTCCTC CCCCTTTCTC TCTAAGGATG 60  
GCCAGAAGG AGAACTCCTA CCCCTGGCCC TACGGCCGAC AGACGGCTCC ATCTGGCCTG 120  
AGCACCTGC CCCAGCGAGT CCTCCGGAAG GAGCCTGTCA CCCCATCTGC ACTTGTCCTC 180  
ATGAGCCGCT CCAATGTCCA GCCACAGCT GCCCTGGCC AGAAGGTGAT GGAGAATAGC 240  
AGTGGGACAC CCGACATCTT AACGCGGCAC TTCACAATTG ATGACTTTGA GATTGGGCGT 300  
CCTCTGGGCA AAGGCAAGTT TGGAAACGTG TACTTGGCTC GGGAGAAGAA AAGCCATTTC 360  
ATCGTGGCGC TCAAGGTCCT CTTCAAGTCC CAGATAGAGA AGGAGGGCGT GGAGCATCAG 420  
CTGCGCAGAG AGATCGAAAT CCAGGCCAC CTGCACCATC CCAACATCCT GCGTCTCTAC 480  
AACTATTTTT ATGACCGGAG GAGGATCTAC TTGATTCTAG AGTATGCCCC CCGCGGGGAG 540  
CTCTACAAGG AGCTGCAGAA GAGCTGCACA TTTGACGAGC AGCGAACAGC CACGATCATG 600  
GAGGAGTTGG CAGATGCTCT AATGTACTGC CATGGGAAGA AGGTGATTCA CAGAGACATA 660  
AAGCCAGAAA ATCTGCTCTT AGGGCTCAAG GGAGAGCTGA AGATTGCTGA CTTGGGCTGG 720  
TCTGTGCATG CGCCCTCCCT GAGGAGGAAG ACAATGTGTG GCACCCTGGA CTACCTGCCC 780  
CCAGAGATGA TTGAGGGGCG CATGCACAAT GAGAAGGTGG ATCTGTGGTG CATTGGAGTG 840  
CTTTGCTATG AGCTGCTGGT GGGGAACCCA CCTTTGAGA GTGCATCACA CAACGAGACC 900  
TATCGCCGCA TCGTCAAGGT GGACCTAAAG TTCCCGGCTT CTGTGCCCAC GGGAGCCCAG 960  
GACCTCATCT CCAAAGTCT CAGGCATAAC CCTCGGAAC GGCTGCCCTT GGCCAGGTC 1020  
TCAGCCACC CTTGGGTCCG GGCCAACCTT CGGAGGGTGC TGCTCCCTC TGCCCTTCAA 1080  
TCTGTGCGCT GATGGTCCCT GTCATTCACT CGGGTGGTGC TGTGTGTATG TCTGTGTATG 1140  
TATAGGGGAA AGAAGGGATC CCTAACTGTT CCCTTATCTG TTTTCTACCT CCTCCTTTGT 1200  
TTAATAAAGG CTGAAGCTTT TTGT

Seq ID NO: 481 Protein sequence  
Protein Accession #: NP\_004208

1 11 21 31 41 51



MAQKENSYPW	PYGRQTAPSG	LSTLPQRVLR	KEPVTPSALV	LMSRSNVQPT	AAPGQKVMEN	60
SSGTPDILTR	HFTIDDFEIG	RPLGKGKFGN	VYLAREKKSH	FIVALKVLFK	SQIEKEGVEH	120
QLRREIEIQA	HLHHPNLR	YNYFYDRRRI	YLILEYAPRG	ELYKELQKSC	TFDEQRTATI	180
MEELADALMY	CHGKKVIHRD	IKPENLLLGL	KGELKIDAFG	WSVHAPSLRR	KTMCGLDYL	240
PPEMIEGRMH	NEKVDLWCIG	VLCYELLVGN	PPFESASHNE	TYRRIVKVDL	KFPASVPTGA	300
QDLISKLLRH	NPSERLPLAQ	VSAHPWVRAN	SRRLPPSAL	QSV		

Seq ID NO: 482 DNA sequence  
Nucleic Acid Accession #: AK055663  
Coding sequence: 38..1423

1	11	21	31	41	51	
AGAACGGCTT	CCGCGGGGAG	CTGTGCAGCT	CCTTATCATG	GGGACAATTC	ATCTCTTTTCG	60
AAAACCACAA	AGATCCTTTT	TTGGCAAGTT	GTTACGGGAA	TTTAGACTTG	TAGCAGCTGA	120
CCGAAGGTCC	TGGAAGATAC	TGCTCTTTGG	TGTAATAAAC	TTGATATGTA	CTGGCTTCCT	180
GCTTATGTGG	TGCAGTTCTA	CTAATAGTAT	AGCTTTAACT	GCCTATACTT	ACCTGACCAT	240
TTTTGATCTT	TTTAGTTTAA	TGACATGTTT	AATAAGTTAC	TGGGTAACAT	TGAGGAAACC	300
TAGCCCTGTC	TATTCATTTG	GGTTTGAAAG	ATTAGAAGTC	CTGGCTGTAT	TTGCCTCCAC	360
AGTCTTGGA	CAGTTGGGAG	CTCTCTTTAT	ATTAAAAGAA	AGTGCAGAAC	GCTTTTTTGA	420
ACAGCCCGAG	ATACACACGG	GAAGATTATT	AGTTGGTACT	TTTGTGGCTC	TTTGTTCAC	480
CCTGTTTACG	ATGCTTTCTA	TTCGGAATAA	ACCTTTTGCT	TATGTCTCAG	AAGCTGCTAG	540
TACGAGCTGG	CTTCAAGAGC	ATGTTGCAGA	TCTTAGTCGA	AGCTTGTGTG	GAATTATTCC	600
GGGACTTAGC	AGTATCTTCC	TTCCCCGAAT	GAATCCATTT	GTTTGTATTG	ATCTTGCTGG	660
AGCATTGCT	CTTTGTATTA	CATATATGCT	CATTGAAATT	AATAATTATT	TTGCCGTAGA	720
CCTGCTCT	GCTATAGCTA	TTGCCTTGAT	GACATTTGGC	ACTATGTATC	CCATGAGTGT	780
GTACAGTGGG	AAAGTCTTAC	TCCAGACAAC	ACCACCCCAT	GTTATTGGTC	AGTTGGACAA	840
ACTCATCAGA	GAGGTATCTA	CCTTAGATGG	AGTTTTAGAA	GTCCGAAATG	AACATTTTTG	900
GACCCTAGGT	TTTGGCTCAT	TGGCTGGATC	AGTGCATGTA	AGAATTCGAC	GAGATGCCAA	960
TGAACAAATG	GTTCTTGCTC	ATGTGACCAA	CAGGCTGTAC	ACTCTAGTGT	CTACTCTAAC	1020
TGTTCAAATT	TTCAAGGATG	ACTGGATTAG	GCCTGCCTTA	TTGTCTGGGC	CTGTTGCAGC	1080
CAATGTCCTA	AACTTTTCAG	ATCATCACGT	AATCCCAATG	CCTCTTTTAA	AGGGTACTGA	1140
TGATTTGAAC	CCAGTTACAT	CAACTCCAGC	TAAACCTAGT	AGTCCACCTC	CAGAATTTTC	1200
ATTTAACACT	CCTGGGAAAA	ATGTGAACCC	AGTTATTCTT	CTAAACACAC	AAACAAGGCC	1260
TTATGGTTTT	GGTCTCAATC	ATGGACACAC	ACCTTACAGC	AGCATGCTTA	ATCAAGGACT	1320
TGGAGTTCCA	GGAATTGGAG	CAACTCAAGG	ATTGAGGACT	GGTTTTACAA	ATATACCAAG	1380
TAGATATGGA	ACTAATAATA	GAATTGGACA	ACCAAGACCA	TGATAGACTC	TAATTATTTT	1440
TTATAAGGAA	TATTGACTCC	TTGGCTTCCA	ATTTATTTAG	TAATCCAAT	TTGCATTGAC	1500
TGTTTAATCA	TTTACTCTAA	ATGTTAGCTA	ATAGTAGTCT	TGTTACATT	TCATGAAACC	1560
TATGAAACTA	TATTTTGTGA	AAATGTATTT	GTGACAGTGA	AATCCTCGTA	AATGTTAAAG	1620
GCTTTAAATA	GGCTTCCTTT	AGAAAATGTG	TTTCTTTAAA	TTTGGATTTT	GGTATCTTTG	1680
GTTTTGTAGT	TGACTGCAGT	GTGATGTGAC	CTTACCTTTA	TAAGAGCCAC	TTGATGGAGT	1740
AGATCTGTCA	CATTACTAAG	ATACGATATT	TCTTTTTTTT	TCCGAGACGG	AGTCTTGCTC	1800
TGCCACTGTG	CCCGGCCAAT	ACATTATTAT	TAAGTTAAGG	CTGTACTTTA	TTAAGGCTTC	1860
CTTAGTTTTT	GTTTTGTTTT	GTTTTTTGAG	ATGGAGTCTC	ACTCTGTGCG	CCAGGCTGGA	1920
ATGCAGTGGC	ATGATCTCAG	CTCACTGCAA	CCTCTGCCTC	CTGAGTTCAA	ATGATTCTCC	1980
TGCCTCAGCC	TCCCGAGTAG	CTGGGATTAC	AGGCACCTGC	CACCACGCCC	AGCTAATTTT	2040
TGTATTTTTA	GTAAAGACGG	GGGATTTTAC	CATGTTGGCC	AGGCTGGTCT	TGAACTCCTG	2100
ACCTCATGAT	CCACCCACCT	TAGCCTCCCA	AAGTGTGGG	ATTAGGTGTG	AGCCACCGCA	2160
CCTGGCCGAT	ATTTCTTTTA	ATGAAATTTA	TAAATATGCT	TCTTGAATAA	TACACATTTT	2220
GGGAAAGGGA	AAAATGTCTG	TTCAAAAAGT	AAAGTCTCT	TTTATAGCTT	TTCCAAACTT	2280
AATTGCTAAA	TTTTCTTTTG	AGGTTCTCCT	GAATTATGTC	TTACAAACTA	AAAGCAAAAA	2340
TTTTTAGCAG	AAATTTTGGA	ATACATTCTA	TCTAGCACAA	TTTGAATTTT	TAATTATCAA	2400
GATTTTTGTT	AAAGTTTCTC	TCCTTTAAAA	ATTTTAGTAC	ATTTGTAAAT		

Seq ID NO: 483 Protein sequence  
Protein Accession #: BAB70980.1

1	11	21	31	41	51	
MGTIHLFRKP	QRSFFGKLLR	EFRLVAADRR	SWKILLFGVI	NLICTGFLLM	WCSSTNSIAL	60
TAYTYLTIFD	LFSMLTCLIS	YWVTLRKPS	VYSFGFERLE	VLAVFASTVL	AQLGALFILK	120
ESAERFLEQP	EIHTGRLLVG	TFVALCFNLF	TMLSIRNKPF	AYVSEAASTS	WLQEHVADLS	180
RSLCGIIPGL	SSIFLPRMNP	FVLIDLAGAF	ALCITYMLIE	INNIFYAVDTA	SAIAIALMTF	240
GTMYPMVSYS	GKVLQTPP	HVIGQLDKLI	REVSTLDGVL	EVREHFVWTL	GFGSLAGSVH	300
VRIRRDANEQ	MVLAHVNTNL	YTLVSTLTQV	IFKDDWIRPA	LLSGPVAANV	LNFSDDHVIP	360
MPLLLKGTDDL	NPVTSTPAKP	SSPPPEFSFN	TPGKNVNPVI	LLNTQTRPYG	FGLNHGHTPY	420
SSMLNQGLGV	PGIGATQGLR	TGFTNIPSR	GTNNRIGQPR	P		

Seq ID NO: 484 DNA sequence  
Nucleic Acid Accession #: FGENSESH predicted  
Coding sequence: 1..900

1	11	21	31	41	51	
ATGCCGCCGC	GGGAGCTGAG	CGAGGCCGAG	CCGCCCCCGC	TCCGGGCCCC	GACCCCTCCC	60
CCGCGGCGGC	GTAGCGCGCC	CCCAGAGCTG	GGCATCAAGT	GCGTGTGCTG	GGGCGACGGC	120
GCCGTGGGCA	AGAGCAGCCT	CATCGTCAGC	TACACCTGCA	ATGGGTACCC	CGCGCGCTAC	180
CGGCCCCACTG	CGCTGGACAC	CTTCTCTGGT	ACGTACGTTT	AATCGCCCGT	GCGGCCGCGT	240
GGCTGCGGCG	GGGCTGTGCA	CCGGGGAGCT	GGGGCGGGCG	TCTCGGCGGG	AGGGCGCAGA	300
GGACCCCGGG	GAGGAGACTG	GAGCAGGCC	CGAGGTGGCG	CTGGTGCAGC	CCAGGACGCT	360
CTTCCTAACT	CAGGCTCTCC	CCGCCCCGCG	CCTGCAGTGC	AAGTCCTGGT	GGATGGAGCT	420
CCGCTGCGCA	TTGAGCTCTG	GGACACAGCG	AGGAGGAGG	ATTTTGACCG	ACTTCGTTCC	480
CTTTGCTACC	CGGATACCGA	TGTCTTCTCG	GCGTGCTTCA	GCGTGGTGCA	GCCCAGCTCC	540
TTTCAAAACA	TCACAGAGAA	ATGGCTGCCC	GAGATCCGCA	CGCACAAACC	CCAGGCGCCT	600
GTGCTGCTGG	TGGGCACCCA	GGCCGACCTG	AGGGACGATG	TCAACGTACT	AATTCAGCTG	660



GACCAGGGGG GCCGGGAGGG CCCCCTGCCC CAACCCAGG CTCAGGGTCT GGCCGAGAAG 720  
 ATCCGAGCCT GCTGCTACCT TGAGTGCTCA GCCTTGACGC AGAAGAACTT GAAGGAAGTA 780  
 TTTGACTCGG CTATTCTCAG TGCCATTGAG CACAAAGCCC GGCTGGAGAA GAAACTGAAT 840  
 GCCAAAGGTG TGCGCACCTT CTCCCGCTGC CGCTGGAAGA AGTTCTTCTG CTTCGTTTGA

Seq ID NO: 485 Protein sequence  
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51  
 | | | | | |  
 MPPRELSEAE PPPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60  
 RPTALDTFSG TYVQSPVRPR GCGGAVHRGA GAGVSAGGRR GPRGGDWSRP RGGAGAAQDA 120  
 LPNSGSPRPA PAVQVLVDGA PVRIELWDTA GQEDFDRLRS LCYPDLDVFL ACFSVVQPSS 180  
 FQNITEKWLP EIRTHNPQAP VLLVGTQADL RDDVNVLQL DQGGREGPVP QPQAQGLAEK 240  
 IRACCYLECS ALTQKNLKEV FDSAILSIE HKARLEKKLN AKGVRTLSRC RWKKFFCFV

Seq ID NO: 486 DNA sequence  
 Nucleic Acid Accession #: XM\_063832.2  
 Coding sequence: 1..711

1 11 21 31 41 51  
 | | | | | |  
 ATGCCGCGGC GGGAGCTGAG CGAGGCCGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC 60  
 CCGCGCGCGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC 120  
 GCCGTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180  
 CGGCCCCACTG CGCTGGACAC CTTCTCTGTG CAAGTCCTGG TGGATGGAGC TCCGGTGCGC 240  
 ATTGAGCTCT GGGACACAGC GGGACAGGAG GATTTTGACC GACTTCGTTT CTTTGCTTAC 300  
 CCGGATACCG ATGTCTTCTT GCGGTGCTTC AGCGTGCTGC AGCCAGCTC CTTTCAAAAC 360  
 ATCACAGAGA AATGGCTGCC CGAGATCCGC ACGCACAAAC CCCAGGCGCC TGTGCTGCTG 420  
 GTGGGCACCC AGGCCGACCT GAGGGACGAT GTCAACGTAC TAATTCAGCT GGACCAGGGG 480  
 GGCCGGGAGG GCCCCGTGCC CCAACCCAG GCTCAGGGTC TGGCCGAGAA GATCCGAGCC 540  
 TGCTGCTACC TTGAGTGCTC AGCCTTGACG CAGAAGAACT TGAAGGAAGT ATTTGACTCG 600  
 GCTATTCTCA GTGCCATTGA GCACAAAGCC CGGCTGGAGA AGAAACTGAA TGCCAAAGGT 660  
 GTGCGCACCC TCTCCCGCTG CCGCTGGAAG AAGTCTTCTT GCTTCGTTTG A

Seq ID NO: 487 Protein sequence  
 Protein Accession #: XP\_063832.1

1 11 21 31 41 51  
 | | | | | |  
 MPPRELSEAE PPPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60  
 RPTALDTFSV QVLVDGAPVR IELWDTAGQE DFDRLRLSCY PDLDVFLACF SVVQPSSFQN 120  
 ITEKWLPEIR THNPQAPVLL VGTQADLRDD VNVLIQLDQG GREGPVPQPQ AQGLAEKIRA 180  
 CCYLECSALT QKNLKEVFD SAILSIEHKA RLEKKLNAG VRTLSRCRWK KFFCFV

Seq ID NO: 488 DNA sequence  
 Nucleic Acid Accession #: NM\_014398.1  
 Coding sequence: 64..1314

1 11 21 31 41 51  
 | | | | | |  
 GGCACCGATT CGGGGCTGTC CCGGACTTCG CCGCACGCTG CAGAACCTCG CCCAGCGCCC 60  
 ACCATGCCCC GCGAGCTCAG CGCGGCGGCC GCGCTCTTCG CGTCCCTGGC CGTAATTTTG 120  
 CACGATGGCA GTCAAATGAG AGCAAAAGCA TTTCCAGAAA CCAGAGATTA TTCTCAACCT 180  
 ACTGCAGCAG CAACAGTACA GGACATAAAA AAACCTGTCC AGCAACCAGC TAAGCAAGCA 240  
 CCTCACCAAA CTTTAGCAGC AAGATTCATG GATGGTCATA TCACCTTTCA AACAGCGGCC 300  
 ACAGTAAAAA TTCCAACAAC TACCCAGCA ACTACAAAAA AACTGCAAC CACCAGCCCCA 360  
 ATTACCTACA CCCTGGTCAC AACCAGGCC ACACCAACA ACTCACACAC AGCTCCTCCA 420  
 GTTACTGAAG TTACAGTCGG CCTAGCTTA GCCCTTAT CACTGCCACC CACCATCACC 480  
 CCACCAGCTC ATACAGCTGG AACCAGTTCA TCAACCGTCA GCCACACAAC TGGGAACACC 540  
 ACTCAACCCA GTAACCAGAC CACCCTTCCA GCAACTTTAT CGATAGCACT GCACAAAAGC 600  
 ACAACCGGTC AGAAGCCTGA TCAACCCACC CATGCCCCAG GAACAACGGC AGCTGCCACC 660  
 AATACCACCC GCACAGCTGC ACCTGCCTCC ACGGTTCTCT GGCCACCCCT TGCACCTCAG 720  
 CCATCGTCAG TCAAGACTGG AATTTATCAG GTTCTAAACG GAAGCAGACT CTGTATAAAA 780  
 GCAGAGATGG GGATACAGCT GATTGTTCAA GACAAGGAGT CGGTTTTTTC ACCTCGGAGA 840  
 TACTTCAACA TCGACCCCAA CGCAAGCAA GCCTCTGGGA ACTGTGGCAC CCGAAAATCC 900  
 AACCTTCTGT TGAATTTTCA GGGCGGATTT GTGAATCTCA CATTTACCAA GGATGAAGAA 960  
 TCATATTATA TCAGTGAAGT GGGAGCCTAT TTGACCGTCT CAGATCCAGA GACAGTTTAC 1020  
 CAAGGAATCA AACATGCGGT GGTGATGTTT CAGACAGCAG TCGGGCATTC CTTCAAGTGC 1080  
 GTGAGTGAAC AGAGCCTCCA GTTGTCAGCC CACCTGCAGG TGAAAACAAC CGATGTCCAA 1140  
 CTTCAAGCCT TTGATTTTGA AGATGACCAC TTTGGAAATG TGGATGAGTG CTCGTCTGAC 1200  
 TACACAATTG TGCTTCCTGT GATTGGGGCC ATCGTGGTTG GTCTCTGCCT TATGGGTATG 1260  
 GGTGTCTATA AAATCCGCCT AAGGTGTCAA TCATCTGGAT ACCAGAGAAT CTAATTGTTG 1320  
 CCCGGGGGGA ATGAAAATAA TGGAATTTAG AGAACTCTTT CATCCCTTCC AGGATGGATG 1380  
 TTGGGAAATT CCTCAGAGT GTGGGTCTTT CAAACAATGT AAACCACCAT CTTCTATTCA 1440  
 AATGAAGTGA GTCATGTGTG ATTTAAGTTC AGGCAGCACA TCAATTTCTA AATACTTTT 1500  
 GTTTATTTTA TGAAAGATAT AGTGAGCTGT TTATTTTCTA GTTTCCTTTA GAATATTTTA 1560  
 GCCACTCAAA GTCAACATTT GAGATATGT GATATACAT AATATATGTA AAGTAGAATA 1620  
 AGCCTTCAAA TTATAAACCA AGGGTCAATT GTAACATAA CTACTGTGTG TGCATTGAAG 1680  
 ATTTTATTTT ACCCTTGATC TTAACAAAGC CTTTGCTTTG TTATCAAATG GACTTTTCA 1740  
 GCTTTTACTA TCTGTGTTT ATGGTTTCAT GTAACATACA TATTCCTGGT GTAGCACTTA 1800  
 ACTCCTTTTC CACTTAAAT TTGTTTTTGT TTTTGTAGAC GGAGTTTCAC TCTTGTCAAC 1860  
 CAGGCTGGAG TACAGTGGCA CGATCTCGG TTAGGCAAC CTCCGCTCC CCGGTTCAAG 1920  
 TGATTCTCCT GTTCAGCTT CCCGATGAG TGGGATTACA GGCACACACT ACCACGCTG 1980  
 GCTAATTTT GTATTTTAT TATAGACGGG TTTCACCATG TTGGCCAGAC TGGTCTTGAA 2040  
 CTCTTGACCT CAGGTGATCC ACCACCTCA GCCTCCCAA GTGCTGGGAT TACAGGCATG 2100  
 AGCCATTGCG CCCGGCCTTA AATGTTTTT TTAATCATCA AAAAGAACAA CATATCTCAG 2160



5  
10  
15

```
GTGTGCTAAG TGTTTTTATG TAAAACCAAC AAAAAGAACA AATCAGCTTA TATTTTTTAT 2220
CTTGATGACT CCTGCTCCAG AATTGCTAGA CTAAGAATTA GGTGGCTACA GATGGTAGAA 2280
CTAAACAATA AGCAAGAGAC AATAATAATG GCCCTTAATT ATTAACAAAG TGCCAGAGTC 2340
TAGGCTAAGC ACTTTATCTA TATCTCATTT CATTCTCACA ACTTATAAGT GAATGAGTAA 2400
ACTGAGACTT AAGGGAAGTG AATCACTTAA ATGTCACCTG GCTAACTGAT GGCAGAGCCA 2460
GAGCTTGAAT TCATGTTGGT CTGACATCAA GGTCTTTGGT CTTCTCCCTA CACCAAGTTA 2520
CCTACAAGAA CAATGACACC ACACTCTGCC TGAAGGCTCA CACCTCATAC CAGCATACGC 2580
TCACCTTACA GGGAAATGGG TTTATCCAGG ATCATGAGAC ATTAGGGTAG ATGAAAGGAG 2640
AGCTTTGCAG ATAACAAAAT AGCCTATCCT TAATAAATCC TCCACTCTCT GGAAGGAGAC 2700
TGAGGGGCTT TGTA AACAT TAGTCAGTTG CTCATTTTAA TGGGATTGCT TAGCTGGGCT 2760
GTAAAGATGA AGGCATCAAA TAAACTCAAA GTATTTTAA ATTTTTTTGA TAATAGAGAA 2820
ACTTCGCTAA CCAACTGTTC TTTCTTGAGT GTATAGCCCC ATCTTGTTGGT AACTTGCTGC 2880
TTCTGCACTT CATATCCATA TTTCTTATG TTCACTTTAT TCTGTAGAGC AGCCTGCCAA 2940
GAATTTTATT TCTGCTGTTT TTTTGTGCTG TAAAGAAAGG AACTAAGTCA GGATGTAAAC 3000
AGAAAAGTCC ACATAACCTT AGAATCTTAA GTCAAGGAAT AATTCAAGTC AGCCTAGAGA 3060
CCATGTTGAC TTTCTCATG TGTTCCTTAA TGACTCAGTA AGTTGGCAAG GTCCTGACTT 3120
TAGTCTTAAT AAAACATTGA ATTGTAGTAA AGGTTTTTGC AATAAAAAC TACTTTGG
```

20 Seq ID NO: 489 Protein sequence  
Protein Accession #: NP\_055213.1

25  
30

```
1 11 21 31 41 51
| | | | |
MPRQLSAAAA LFASLAVILH DGSQMRKAF PETRDYSQPT AAATVQDIKK PVQQPAKQAP 60
HQTAAARFMD GHITFQTAAT VKIPTTTPAT TKNTATTSPI TYTLVTTQAT PNNSHTAPPV 120
TEVTVGPSLA PYSLPPTITP PAHTAGTSSS TVSHTTGNTT QPSNQTTLPA TLSIALHKST 180
TGQKPDQPTH APGTTAAAHN TTRTAAPAST VPGPTLAPQP SSVKTGIYQV LNGSRLCIKA 240
EMGIQLIVQD KESVFSPPRY FNIDPNATQA SGNCGRKSN LLLNFQGGFV NLFTKDEES 300
YYISEVGAYL TVSDPETVYQ GIKHAVVMFQ TAVGHSFKCV SEQSLQLSAH LQVKTTDVQL 360
QAFDFEDDHF GNVDECSSDY TIVLPVIGAI VVGLCLMGMG VYKIRLRCQS SGYQRI
```

35 Seq ID NO: 490 DNA sequence  
Nucleic Acid Accession #: NM\_005409.3  
Coding sequence: 94..378

40  
45  
50  
55  
60

```
1 11 21 31 41 51
| | | | |
TTCCTTTCAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAAGCT GAAGTAGCAG 60
CAACAGCACC AGCAGCAACA GCAAAAAACA AACATGAGTG TGAAGGGCAT GGCTATAGCC 120
TTGGCTGTGA TATTGTGTGC TACAGTTGTT CAAGGCTTCC CCATGTTCAA AAGAGGACGC 180
TGTCTTTGCA TAGGCCCTGG GGTAAAAGCA GTGAAAGTGG CAGATATTGA GAAAGCCTCC 240
ATAATGTACC CAAGTAACAA CTGTGACAAA ATAGAAGTGA TTATTACCCT GAAAGAAAAAT 300
AAAGGACAAC GATGCCTAAA TCCCAAATCG AAGCAAGCAA GGCTTATAAT CAAAAAAGTT 360
GAAAGAAAGA ATTTTAAAAA ATATCAAAAC ATATGAAGTC CTGGAAAAGG GCATCTGAAA 420
AACCTAGAAC AAGTTTAACT GTGACTACTG AAATGACAAG AATTCTACAG TAGGAAACTG 480
AGACTTTTCT ATGGTTTTGT GACTTTC AAC TTTGTACAG TTATGTGAAG GATGAAAGGT 540
GGGTGAAAGG ACCAAAAACA GAAATACAGT CTTCTGAAT GAATGACAAT CAGAATTCCA 600
CTGCCCAAAG GAGTCCAGCA ATTAAATGGA TTTCTAGGAA AAGCTACCTT AAGAAAGGCT 660
GGTTACCATC GGAGTTTACA AAGTGCTTTC ACGTTCTTAC TTGTTGTATT ATACATTCAT 720
GCATTTCTAG GCTAGAGAAC CTTCTAGATT TGATGCTTAC AACTATTCTG TTGTGACTAT 780
GAGAACATTT CTGTCTCTAG AAGTTATCTG TCTGTATTGA TCTTTATGCT ATATTACTAT 840
CTGTGGTTAC AGTGGAGACA TTGACATTAT TACTGGAGTC AAGCCCTTAT AAGTCAAAAG 900
CATCTATGTG TCGTAAAGCA TTCTTCAAC ATTTTTCAT GCAAATACAC ACTTCTTTCC 960
CCAAATATCA TGTAGCACAT CAATATGTAG GGAACATTC TTATGCATCA TTTGGTTTGT 1020
TTTATAACCA ATTCATTAAA TGTAATTCAT AAAATGTACT ATGAAAAAAA TTATACGCTA 1080
TGGGATACTG GCAACAGTGC ACATATTTCA TAACCAAAAT AGCAGCACCG GTCTTAATTT 1140
GATGTTTTTC AACTTTTATT CATTGAGATG TTTTGAAGCA ATTAGGATAT GTGTGTTTAC 1200
TGTACTTTTT GTTTTGATCC GTTTGTATAA ATGATAGCAA TATCTTGGAC ACATTTGAAA 1260
TACAAAATGT TTTTGTCTAC CAAAGAAAAA TGTGAAAAA TAAGCAAATG TATACCTAGC 1320
AATCACTTTT ACTTTTTGTA ATTCTGTCTC TTAGAAAAAT ACATAATCTA ATCAATTCT 1380
TTGTTTCATG CTATATACTG TAAAATTTAG GTATACTCAA GACTAGTTTA AAGAATCAAA 1440
GTCATTTTTT TCTCTAATAA ACTACCACAA CCTTCTTTT TTAACAAAAA AAA
```

65 Seq ID NO: 491 Protein sequence  
Protein Accession #: NP\_005400.1

70

```
1 11 21 31 41 51
| | | | |
MSVKGMAIAL AVILCATVVQ GFPMFKRGRC LCIGPGVKAV KVADIEKASI MYPSNNCDKI 60
EVIITLKENK GQRCLNPKSK QARLIKKVE RKNF
```

75 Seq ID NO: 492 DNA sequence  
Nucleic Acid Accession #: NM\_000577.1  
Coding sequence: 41..520

80  
85

```
1 11 21 31 41 51
| | | | |
GGCACGAGGG GAAGACCTCC TGTCCATCA GGCCCTCCCC ATGGCTTTAG AGACGATCTG 60
CCGACCCTCT GGGAGAAAAT CCAGCAAGAT GCAAGCCTTC AGAATCTGGG ATGTTAACCA 120
GAAGACCTTC TATCTGAGGA ACAACCAACT AGTTGCCGGA TACTTGCAAG GACCAAAATG 180
CAATTTAGAA GAAAAGATAG ATGTGGTACC CATTGAGCCT CATGCTCTGT TCTTGGGAAT 240
CCATGGAGGG AAGATGTGCC TGTCTGTGT CAAGTCTGGT GATGAGACCA GACTCCAGCT 300
GGAGGCAGTT AACATCACTG ACCTGACGGA GAACAGAAAG CAGGACAAGC GCTTCGCCTT 360
CATCCGCTCA GACAGTGGCC CCACCACCAG TTTTGAAGT GCGCCTGCC CCGGTTGGTT 420
CCTCTGCACA GCGATGGAAG CTGACCAGCC CGTCAGCCTC ACCAATATGC CTGACGAAGG 480
CGTCATGGTC ACCAAATTCT ACTTCCAGGA GGACGAGTAG TACTGCCAGG GCCTGCCTGT 540
TCCCATTTCT GCATGGCAAG GACTGCAGGG ACTGCCAGTC CCCCTGCCCC AGGGCTCCCC 600
```



5	GCTATGGGGG	CACTGAGGAC	CAGCCATTGA	GGGGTGGACC	CTCAGAAGGC	GTCACAACAA	660
	CCTGGTCACA	GGACTCTGCC	TCCTCTTCAA	CTGACCAGCC	TCCATGCTGC	CTCCAGAATG	720
	GTCTTTCTAA	TGTGTGAATC	AGAGCACAGC	AGCCCCTGCA	CAAAGCCCTT	CCATGTCGCC	780
	TCTGCATTCA	GGATCAAACC	CCGACCACCT	GCCCAACCTG	CTCTCCTCTT	GCCACTGCCT	840
	CTTCCTCCCT	CATTCCACCT	TCCCATGCCC	TGGATCCATC	AGGCCACTTG	ATGACCCCCA	900
10	ACCAAGTGGC	TCCCACACCC	TGTTTTACAA	AAAAGAAAAG	ACCAGTCCAT	GAGGGAGGTT	960
	TTTAAGGGTT	TGTGGAAAAT	GAAAATTAGG	ATTTTCATGAT	TTTTTTTTTT	CAGTCCCCGT	1020
	GAAGGAGAGC	CCTTCATTG	GAGATTATGT	TCTTTCGGGG	AGAGGCTGAG	GACTTAAAAT	1080
	ATTCCTGCAT	TTGTGAAAAT	ATGGTGAAAG	TAAGTGGTAG	CTTTTCCCTT	CTTTTTCTTC	1140
	TTTTTTTGTG	ATGTCCCAAC	TTGTAAAAAT	TAAAAGTTAT	GGTACTATGT	TAGCCCCATA	1200
15	ATTTTTTTTT	TCCTTTTAAA	ACACTTCCAT	AATCTGGACT	CCTCTGTCCA	GGCACTGCTG	1260
	CCCAGCCTCC	AAGCTCCATC	TCCACTCCAG	ATTTTTTACA	GCTGCCTGCA	GTACTTTACC	1320
	TCCTATCAGA	AGTTTCTCAG	CTCCCAAGGC	TCTGAGCAAA	TGTGGCTCCT	GGGGGTCTCT	1380
	TCTTCCTCTG	CTGAAGGAAT	AAATTGCTCC	TTGACATTGT	AGAGCTTCTG	GCACTTGAG	1440
	ACTTGATGA	AAGATGGCTG	TGCCTCTGCC	TGTCTCCCCC	ACCAGGCTGG	GAGCTCTGCA	1500
20	GAGCAGGAAA	CATGACTCGT	ATATGTCTCA	GGTCCCTGCA	GGGCCAAGCA	CCTAGCCTCG	1560
	CTCTTGCGAG	GTACTCAGCG	AATGAATGCT	GTATATGTTG	GGTGCAAAGT	TCCCTACTTC	1620
	CTGTGACTTC	AGCTCTGTTT	TACAAATAAA	TCTTGAAAAT	GCCTAAAAAA	AAAAAAAAAA	1680
	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAA			

Seq ID NO: 493 Protein sequence  
Protein Accession #: NP\_000568.1

25	1	11	21	31	41	51	
	MALETICRPS	GRKSSKMQAF	RIWDVNQKTF	YLRNNQLVAG	YLQGPVNLE	EKIDVVPFIEP	60
	HALFLGIHGG	KMCLSCVKSG	DETRLQLEAV	NITDLSENKR	QDKRFAFIRS	DSGPTTSFES	120
	AACPGWFLCT	AMEADQPVSL	TNMPDEGVMV	TKFYFQEDE			

30 Seq ID NO: 494 DNA sequence  
Nucleic Acid Accession #: NM\_002081.1  
Coding sequence: 222..1898

35	1	11	21	31	41	51	
	GGCTGCCCCG	GCGAGCGTTC	GGACCTCGCA	CCCCGCGCGC	CCCGCGCCGC	CGCCGCGGCC	60
40	GGCTTTTGTT	GTCTCCGCCT	CCTCGGCCCG	CGCCGCCTCT	GGACCGCGAG	CCGCGCGCGC	120
	CGGGACCTTG	GCTCTGCCCT	TCGCGGGCGG	GAACTGCGCA	GGACCCGGCC	AGGATCCGAG	180
	AGAGGCGCGG	GCGGGTGGCC	GGGGGCGCCG	CCGGCCCCCG	CATGGAGCTC	CGGGCCCGAG	240
45	GCTGCTGGCT	GCTATGTGCG	GCCGACGCGC	TGGTCTGCTG	CGCCCGCGGG	GACCCGGCCA	300
	GCAAGAGCCG	GAGCTGCGGC	GAGGTCCGCC	AGATCTACGG	AGCCAAGGGC	TTCAGCCTGA	360
	GCGACGTGCC	CCAGGCGGAG	ATCTCGGGTG	AGCACCTGCG	GATCTGTCCC	CAGGGCTACA	420
50	CCTGCTGCAC	CAGCGAGATG	GAGGAGAACC	TGGCCAACCG	CAGCCATGCC	GAGCTGGAGA	480
	CCGCGCTCCG	GGACAGCAGC	CGCTCCCTGC	AGGCCATGCT	TGCCACCCAG	CTGCGCAGCT	540
	TCGATGACCA	CTTCCAGCAC	CTGCTGAACG	ACTCGGAGCG	GACGCTGCAG	GCCACCTTCC	600
55	CCGGCGCCTT	CGGAGAGCTG	TACACGCAGA	ACGCGAGGGC	CTTCCGGGAC	CTGTACTCAG	660
	AGCTGCGCCT	GTACTACCGC	GGTGCCAACC	TGCACCTGGA	GGAGACGCTG	GCCGAGTTCT	720
	GGGCCCCCCT	GCTCGAGCGC	CTCTTCAAGC	AGCTGCACCC	CCAGCTGCTG	CTGCCTGATG	780
60	ACTACCTGGA	CTGCCTGGGC	AAGCAGGCCG	AGGCGCTGCG	GCCCTTCGGG	GAGGCCCCGA	840
	GAGAGCTGCG	CCTGCGGGCC	ACCCGTGCCT	TCGTGGCTGC	TCGCTCCTTT	GTGCAGGGCC	900
	TGGGCGTGCC	CAGCGACGTG	GTCCGGAAG	TGGCTCAGGT	CCCCCTGGGC	CCGGAGTGCT	960
65	CGAGAGCTGT	CTGAAGCTG	GTCTACTGTG	CTCACTGCCT	GGGAGTCCCC	GGCGCCAGGC	1020
	CCTGCCCTGA	CTATTGCCGA	AATGTGCTCA	AGGCTGCTCT	TGCCAACCAG	GCCGACCTGG	1080
	ACGCCGAGTG	GAGGAACCTC	CTGGACTCCA	TGGTGTCTAT	CACCGACAAG	TTCTGGGGTA	1140
70	CATCGGGTGT	GGAGAGTGTC	ATCGGCAGCG	TGCACACGTG	GCTGGCGGAG	GCCATCAACG	1200
	CCCTCCAGGA	CAACAGGGAC	ACGCTCACGG	CCAAGGTCAT	CCAGGGCTGC	GGGAACCCCA	1260
	AGGTCAACCC	CCAGGGCCCT	GGGCCTGAGG	AGAAGCGGCG	CCGGGGCAAG	CTGGCCCCGC	1320
75	GGGAGAGGCC	ACCTTCAGGC	ACGCTGGAGA	AGCTGGTCTC	TGAAGCCAAG	GCCCAGCTCC	1380
	GCGAGCTCCA	GGACTTCTGG	ATCAGCCTCC	CAGGGACACT	GTGCAGTGAG	AAGATGGCCC	1440
	TGAGCACTGC	CAGTGATGAC	CGCTGTGGA	ACGGGATGGC	CAGAGGCCGG	TACCTCCCCG	1500
80	AGGTCATGGG	TGACGGCCTG	GCCAACCAGA	TCAACAACCC	CGAGGTGGAG	GTGGACATCA	1560
	CCAAGCCGGA	CATGACCATC	CGGCAGCAGA	TCATGCAGCT	GAAGATCATG	ACCAACCGGC	1620
	TGCGCAGCGC	CTACAACGGC	AACGACGTGG	ACTTCCAGGA	CGCCAGTGAC	GACGGCAGCG	1680
85	GCTCGGGCAG	CGGTGATGGC	TGCTGGATG	ACCTCTGCGG	CCGGAAGGTC	AGCAGGAAGA	1740
	GCTCCAGCTC	CCGGACGCC	TTGACCCATG	CCCTCCAGG	CCTGTGAGAG	CAGGAAGGAC	1800
	AGAAGACCTC	GGCTGCCAGC	TGCCCCCAGC	CCCCGACCTT	CCTCCTGCCC	CTCCTCCTCT	1860
90	TCCTGGCCCT	TACAGTAGCC	AGGCCCCGGT	GGCGGTAACT	GCCCCAAGGC	CCCAGGGACA	1920
	GAGGCCAAGG	ACTGACTTTG	CCAAAAATAC	AACACAGACG	ATATTTAATT	CACCTCAGCC	1980
	TGGAGAGGCC	TGGGGTGGGA	CAGGGAGGGC	CGGCGGCTCT	GAGCAGGGGC	AGGCGCAGAG	2040
95	GTCCAGGCC	CAGGCCTGGC	CTCGCCTGCC	TTTCTGCCTT	TTAATTTTGT	ATGAGGTCCT	2100
	CAGGTGAGCT	GGGAGCCAGT	GTGCCCAAAA	GCCATGTATT	TCAGGGACCT	CAGGGGCACC	2160
	TCCGGCTGCC	TAGCCCTCCC	CCCAGCTCCC	TGCACCGCCG	CAGAAGCAGC	CCCTCGAGGC	2220
100	CTACAGAGGA	GGCCTCAAAG	CAACCCGCTG	GAGCCACAG	CGAGCCTGTG	CCTTCTCTCC	2280
	CGCCTCTCTC	CACTGGGACT	CCCAGCAGAG	CCCACCAGCC	AGCCCTGGCC	CACCCCCCAG	2340
	CCTCCAGAGA	AGCCCCGCAC	GGGCTGTCTG	GGTGTCCGCC	ATCCAGGGTC	TGGCAGAGCC	2400
105	TCTGAGATGA	TGCATGATGC	CCTCCCCCTC	GCGCAGGCTG	CAGAGCCCCG	CCCCACCTCC	2460
	CTGCGCCCTT	GAGGGGCCCC	AGCGTCTGCA	GGGTGACGCC	TGAGACAGCA	CCACTGCTGA	2520
	GGAGTCTGAG	GACTGTCTTC	CCACGACCCC	TGCAGTGAGG	GGCCCTCCAT	GCGCAGATGA	2580
110	GGGGCCACTG	ACCCACCTGC	GCTTCTGCTG	GAGGAGGGGA	AGCTGGGCCC	AAAGGCCCCAG	2640
	GGAGGCAGCG	TGGGCTCTGC	CAATGTGGGC	TGCCCCCTCG	ACACAGGGCT	CACAGGGCAG	2700
	GCCTTGCTGG	GGTCCAGGGC	TGTTGGAGGA	CCCCGAGGGC	TGAGGAGCAG	CCAGGACCCG	2760
115	CCTGCTCCCA	TCCTCACCCA	GATCAGGAAC	CAGGGCCTCC	CTGTTACCGG	TGACACAGGT	2820
	CAGGGCTCAG	AGTGACCTTC	GGCTGTCAAC	TGCTTACAGG	GATGCTGGTG	GCTGGTGAGA	2880
	CCCGCAGCTG	CACACGGGAA	TGCCTAGTTC	CCTTCCCGAC	CCAGCCAGCT	GCACTGCAGG	2940
120	GCACGGGGAC	CTGGATAGTT	AAGGGCTTTT	CCAAACATGC	ATCCATTTAC	TGACACTTCC	3000
	TGTCCTTGTT	CATGGAGAGC	TGTTGCTGCC	TCCAGATGG	CTTCGGAGGC	CCGCGGGGCC	3060
	CACCTTGGAC	CCTGGTGACC	TCCTGTCACT	CACGTAGGCC	ATCAGGGCCC	TGCCCCAGGC	3120



CTGGACGGGC CCTCCTTCCC TCCTGTGCCC CAGCTGCCAG GTGGCCCTGG GGAGGGGTGG 3180  
 TGTGGTGTGG GGAAGGGGTC CTGCAGGGGG AGGAGGACTT GGAGGGTCTG GGGGCAGCTG 3240  
 TCCTGAACCG ACTGACCCTG AGGAGGCCCG TTAGTGCTGC TTTGCTTTTC ATCACCGTCC 3300  
 CGCACAGTGG ACGGAGGTCC CCGTTGCTG GTCAGGTCCC CATGGCTTGT TCTCTGGAAC 3360  
 CTGACTTTAG ATGTTTTTGGG ATCAGGAGCC CCCAACACAG GCAAGTCCAC CCCATAATAA 3420  
 CCCTGCCAGT GCCAGGGTGG GCTGGGGACT CTGGCACAGT GATGCCGGGC GCCAGGACAG 3480  
 CAGCACTCCC GCTGCACACA GACGGCCTAG GGGTGGCGCT CAGACCCAC CCTACGCTCA 3540  
 TCTCTGGAAG GGGCAGCCCT GAGTGGTCAC TGGTCAGGGC AGTGGCCAAG CCTGCTGTGT 3600  
 CCTTCCTCCA CAAGGTCCCC CCACCGCTCA GTGTACGCGG GTGACGTGTG TTCTTTTGAG 3660  
 TCCTTGATG AATAAAAGGC TGGAAACCTA AA

Seq ID NO: 495 Protein sequence  
 Protein Accession #: NP\_002072.1

1 11 21 31 41 51  
 | | | | |  
 MELRARGWWL LCAAAALVAC ARGDPASKSR SCGEVRQIYG AKGFSLSQVQ QAEISGEHLR 60  
 ICPQGYTCCT SEMEENLANR SHAELETALR DSSRVLQAML ATQLRSFDDH FOHLNDLNDSE 120  
 TLQATFPFAGF GELYTQNLAR FRDLYSELRL YLRGANLHLE ETLAEFWARL LERLFKQLHP 180  
 QLLLPDDYLD CLGKQAEALR PFGGAPRELRL LRATRAFVAA RSFVQGLGVA SDVVRKVAQV 240  
 PLGPECSRAV MKLVYCAHCL GVPGARPCPD YCRNVKLGCL ANQADLDAEW RNLLDSMVL 300  
 TDKFWGTSGV ESIVGSVHTW LAEAINALQD NRDTLTAKVI QGCGNPKVNP QGPGPEEKRR 360  
 RGKLAPEVRP PSGLTLEKLV EAKAQLRDVQ DFWISLPGL CSEKMALSTA SDDRCWNGMA 420  
 RGRYLPEVMG DGLANQINNP EVEVDITKPD MTIRQQIMQL KIMTNRLRSA YNGNDVDFQD 480  
 ASDDGSGSGS

Seq ID NO: 496 DNA sequence  
 Nucleic Acid Accession #: NM\_001650.2  
 Coding sequence: 40.1011

1 11 21 31 41 51  
 | | | | |  
 GGGGCAGGCA ATGAGAGCTG CACTCTGGCT GGGGAAGGCA TGAGTGACAG ACCCACAGCA 60  
 AGGCGGTGGG GTAAGTGTGG ACCTTTGTGT ACCAGAGAGA ACATCATGGT GGCTTTCAAA 120  
 GGGGTCTGGA CTCAAGCTTT CTGGAAGGCA GTCACAGCGG AATTTCTGGC CATGCTTATT 180  
 TTTGTTCTCC TCAGCCTGGG ATCCACCATC AACTGGGGTG GAACAGAAAA GCCTTTACCG 240  
 GTCGACATGG TTCTCATCTC CCTTTGCTTT GGAATCAGCA TTGCAACCAT GGTGCAGTGC 300  
 TTTGGCCATA TCAGCGGTGG CCACATCAAC CTGTCAGTGA CTGTGGCCAT GGTGTGCACC 360  
 AGGAAGATCA GCATCGCCAA GTCTGTCTTC TACATCGCAG CCCAGTGCCT GGGGGCCATC 420  
 ATTGGAGCAG GAATCCTCTA TCTGGTCACA CCTCCAGTGG TGGTGGGAGG CCTGGGAGTC 480  
 ACCATGGTTC ATGGAATCTT TACCGCTGGT CATGGTCTCC TGGTTGAGTT GATAATCACA 540  
 TTTCAATTGG TGTCTACTAT CTTTGCCAGC TGTGATTCCA AACGGACTGA TGTCAGTGGC 600  
 TCAATAGCTT TAGCAATTGG ATTTCTGTT GCAATTGGAC ATTTATTTGC AATCAATTAT 660  
 ACTGGTGCCA GCATGAATCC CGCCGATCC TTTGGACCTG CAGTTATCAT GGGAAATTGG 720  
 GAAACCAATT TAGATATATT GGTGGGCCCC ATCATAGGAG CTGTCTCTCG TGGTGGCCTT 780  
 TATGAGTATG TCTTCTGTCC AGATGTTGAA TTCAAACGTC GTTTTAAAGA AGCCTTCAGC 840  
 AAAGCTGCCC AGCAAACAAA AGGAAGCTAC ATGGAGGTGG AGGACAACAG GAGTCAGGTA 900  
 GAGACGGATG ACCTGATTCT AAAACCTGGA GTGGTGCATG TGATTGACGT TGACCGGGGA 960  
 GAGGAGAAGA AGGGGAAAGA CCAATCTGGA GAGGTATTGT CTTCAAGTATG ACTAGAAGAT 1020  
 CGCACTGAAA GCAGACAAGA CTCCTTAGAA CTGTCTCTAG ATTTCTCTCC ACCCATTAAG 1080  
 GAAACAGATT TGTATATAAT TAGAAATGTG CAGGTTTGTG GTTTCATGTC ATATTACTCA 1140  
 GTCTAAACAA TAAATATTTC ATAATTACAA AAGGAGGAAC GGAAGAAACC TATTGTGAAT 1200  
 TCCAAATCTA AAAAAAGAAA TATTTTAAAG ATGTTCTTAA GCAAATATAT ACCTATTTTA 1260  
 TCTAGTTACC TTTCATTAAC AACCATTAT TAAACGTGTG CAAGATTGTT TTAAGTCTTG 1320  
 CCTGACAGAA CTCAAAGACA CGTCTATCAG CTTATTCCTT CTCTACTGGA ATATTGGTAT 1380  
 AGTCAATTCT TATTGAATA TTTATCTAT TAAACTGAGT TTAACAATGG C

Seq ID NO: 497 Protein sequence  
 Protein Accession #: NP\_001641.1

1 11 21 31 41 51  
 | | | | |  
 MSDRPTARRW GKCGPLCTRE NIMVAFKGVW TQAFWKAVTA EFLAMLIFVL LSLGSTINWG 60  
 GTEKPLPVDV VLISLCFGLS IATMVQCFGH ISGGHINPAV TVAMVCTRKI SIAKSVFYIA 120  
 AQCLGAIIGA GILYLVTPPS VVGGLGVTMV HGNLTAGHGL LVELIITFQL VFTIFASCDS 180  
 KRTDVTGSIA LAIGFSVAIG HLFAINVTGA SMNPARSFGP AVIMGNWENH WIYWVGPIIG 240  
 AVLAGGLYEV VFCPDVEFKR RFKEAFSKAA QQTGKSYMEV EDNRSQVETD DLILKPGVVH 300  
 VIDVDRGEEK KGKDQSGEVL SSV

Seq ID NO: 498 DNA sequence  
 Nucleic Acid Accession #: AB020684.1  
 Coding sequence: 1..1744

1 11 21 31 41 51  
 | | | | |  
 CCCCCCTGTC ATTAATACAT TAAAAAGATT CAATCTTTAC CCTGAGGTAA TTTTGGCCAG 60  
 TTGGTACCGG ATTTATACCA AAATAATGGA CTTGATTGGT ATTCAAACCA AGATATGTTG 120  
 GACGGTTACC AGAGGAGAAG GACTCAGTCC TATTGAAAGC TGTGAAGGAT TGGGAGATCC 180  
 TGCTTGCTTT TATGTTGCTG TAATTTTAT TTTAAATGGA CTAATGATGG CATTATTCTT 240  
 CATATATGGC ACATATTTAA GTGGCAGCCG ATTAGGAGGC CTGGTTACAG TGTGTGCTT 300  
 CTTTTTCAAT CATGGAGAGT GTACCCGTGT AATGTGGACA CCACCTCTCC GTGAAAGCTT 360  
 CTCATATCCA TTTCTTGTTT TTCAGATGTT GCTAGTGACT CATATTCTCA GGGCTACAAA 420  
 ACTTTATAGA GGAAGCTTGA TTGCACTCTG CATTTCCAAT GTATTTTTC TGTCTCCTTG 480  
 GCAGTTTGCT CAGTTTGTAC TTCTTACTCA GATTGCATCA TTATTGTCAG TATATGTTGT 540  
 CGGGTACATT GATATATGTA AATTACGGAA GATCATTTAT ATACACATGA TTTCTCTTGC 600  
 ACTTTGTTTT GTTTTGATGT TTGGGAATCT AATGTTATTA ACTTCTTATT ATGCTTCTTC 660  
 TTTGGTAATT ATTTGGGTA TTCTGGCAAT GAAACCACAT TTCCTGAAA TAAATGTATC 720



	TGAACCTAGT	TTATGGGTTA	TTCAAGGATG	TTTTTGGTTA	TTTGGAAGCTG	TCATACTTAA	780
	ATACTTGACA	TCTAAAAATT	TTGGTATTGC	AGATGACGCT	CATATTGGCA	ACTTACTAAC	840
	ATCAAAATTC	TTTAGTTATA	AGGATTTTGA	TACTTTATTG	TATACCTGTG	CAGCGGAGTT	900
5	TGACTTTATG	GAAAAAGAGA	CTCCACTGAG	ATACACAAAG	ACATTATTGC	TTCCAGTTGT	960
	TCTTGTAGTG	TTTGTGCTA	TTGTTAGAAA	GATTATTAGT	GATATGTGGG	GTGTCTTAGC	1020
	TAAACAACAG	ACACATGTAA	GAAAAACACCA	GTTTGATCAT	GGAGAGCTGG	TTTACCATGC	1080
	ATTGCAATTG	TTAGCATATA	CAGCCCTTGG	TATTTTAATT	ATGAGACTAA	AACTCTTCTT	1140
	GACACCACAC	ATGTGTGTTA	TGGCATCACT	GATCTGCTCA	AGACAGCTAT	TTGGATGGCT	1200
10	CTTTTGCAAA	GTACATCCTG	GTGCTATTGT	GTTTGCTATA	TTAGCAGCAA	TGTCAATACA	1260
	AGGTTGAGCA	AATCTGCAAA	CCCAGTGGAA	TATTGTAGGG	GAGTTCAGCA	ATTTGCCCCA	1320
	AGAAGAACTT	ATAGAATGGA	TCAAATATAG	TACTAAACCA	GATGCAGTGT	TTGCGGGTGC	1380
	CATGCCCACG	ATGGCAAGTG	TTAAGCTCTC	TGCACTTCGG	CCCATTGTGA	ATCATCCACA	1440
	TTATGAAGAC	GCAGGCTTGA	GAGCCAGAAC	AAAAATAGTA	TACTCAATGT	ATAGTCGGAA	1500
15	AGCAGCCGAA	GAAGTGAAGC	GAGAACTGAT	AAAGTTAAAA	GTGAACTATT	ACATTCTAGA	1560
	AGAGTCATGG	TGTGTAAGAA	GATCCAAGCC	TGGTTGCAGT	ATGCCTGAAA	TTTGGGATGT	1620
	AGAAGATCCT	GCCAAATGCTG	GGAAAACTCC	CTTATGTAAC	CTCTTGGTGA	AGGATTCCAA	1680
	ACCTCACTTC	ACCACTGTAT	TCCAGAACAG	TGTTTACAAA	GTCCTAGAAG	TTGTAAAAGA	1740
	ATGACTGCTA	CATGACCTGC	TGCCTACGGA	GAACCTACATC	TGTAATGGTT	TTAATGTTTT	1800
20	GCTAAGTCAT	GTGTTGTTCA	TATCCCAAAA	ACTTTTATAG	GTAACGTGTT	TCAAATAGAA	1860
	AACGTTTTAT	TTGGTCAATT	TGAATGTCAT	TCTAATTATA	AAAATGACTT	ACACCTTTAT	1920
	CAATTGGTTA	CTATTTCAAT	GCACCTTTTA	AAATTTGCTA	TGCAAATGAG	TATATGCTTG	1980
	TACTTGACTT	TAATATTTGT	GCTAAAGTGA	GCAAAGCTAC	CTGTATAAAG	AAAACACAGT	2040
	GGGTTGTGAC	AAGGATGACA	TGAAAATACA	GGACAATTCT	GACAATGTAG	GGGCTGATTT	2100
25	TATAGTGTA	GAACATTAA	TGCCCTTGC	TTCTTTTTTC	TGCCTCTTGC	TCTTGTCTTT	2160
	TGGACATTTT	AGTGATTGTA	AGTTCTTCGG	TCATGTCAGC	CCCTGTCATC	AACTTGAGTT	2220
	ACAGTAGATG	GGGCAGACAT	GGAGTGTGTT	CTATATAAAA	CTATCTGTTT	GTTTTACTTC	2280
	CTTGTCGCGT	TTTTGTTCTC	TGTTCTCTTG	TTATCTGAAGC	TTTCTCTGCC	CATTATTAAT	2340
	CCAAACTCTT	GGACCTTGTG	GTTAGGAAAT	TCCCTTAACT	TCCAGCCATA	TGGCATTATC	2400
30	GTGTCTCTTT	CTCTCTCTCT	CTTGCTCTCT	CTCTTCTCCT	CTTCCCCATA	TTTTCTGTCA	2460
	AATAAGTACT	GTTTACTCAT	TTAGTTGCTT	ATCAAGTACT	TATTCTTGGT	TTTAAAAAAA	2520
	ATTAATGGTA	ACTGTATTTT	TCTCATTTTT	AGCATTATTC	AAATGTTTAT	ATTTTAATAC	2580
	CTTTAAACCA	CTTTAAAGTT	TTTTCATGTT	TAATTATAGT	TTTAAGAAAA	ACTATTTTGA	2640
	ACAACCCCAA	ATATAGTGCA	TCTAGAAACT	AATGTATATT	TGATTAGACA	TCATTTATAG	2700
35	TGGAACAGTA	GACTGTAGTA	CATGGTAATT	TTTCTTTTAC	TATTAAGATA	CAATAAAACA	2760
	TGACTAATTT	TGCTGTCAAA	AATGTAAAGA	ATAATGATAA	ATGGAGTTTT	TTATATTTTA	2820
	CTTTTAAGAT	TGCCTGTCTT	TAATAAGACA	AAGCCTTAAG	CCTTATGTTA	TAATTTTGGT	2880
	TCTAAAAACC	ATCATTTTCA	TATAAGGAAT	AAGTATATTT	CGTCTCCTC	TTTAGTTTTT	2940
	TTCTTCTTAT	TTATTTTAT	TTTGAAAAAT	TTCTACACCT	TCTTTGAATT	CCTTGTATGA	3000
40	ATTTTGTGTT	CTTAGAAGTT	AATTTGTGTG	AAATGAGATT	CTTCAAAACG	ATGAAACCTC	3060
	ATAGCTCTGA	GAAAAGGTTT	TAGGGTTTTA	AAATCTAAGC	AAAGCGTGAC	TATGGCTGAC	3120
	AGACTACACA	TTTAATTATA	CAGCTTCTCT	TTCTTAACCA	CAGGCAGATT	AACCTCATTG	3180
	TGGATTGTCC	TTTACACCTT	AGTCCTCAGG	CATGGTTTCT	GGTGCCCACT	CCTGGAAGCC	3240
	GCTGTTCCCT	TTTACCTTTC	TTACCAGAGC	CCAAGGGCAG	GCCTGGTCCC	GGGGAAGCAG	3300
45	CAGCTTGCTG	ACATAAGTCA	GCTGCAAAAG	CTGAGGAGTG	TGCCCTCAGA	GAAGCACCGC	3360
	CCCCAGTCT	TGTGCCAGCG	CCTAGAGCTC	CAGCTCCCAG	GGATGCTCCT	TCCCTGGAGG	3420
	CAGCCCAGGA	GAGGGACTCT	GGCAGCGTTC	TTTCAATTTG	TGGCCACTGT	TTCTCATTG	3480
	CTGGTTGACT	GTTTTTATTT	CTTAGGCTTT	TGCTAGTTTT	AGAAAATAGG	GAAGCAGCCC	3540
	TTGATTTGTG	GATTTAAAGC	AACATTTGAG	CGATGATGCA	CAACAGTCCA	GGAAAATGGG	3600
50	CGGTGGACAC	TTGAGGCTGA	GGATGGGAGT	TGACATGAGC	AGGGAGAGGG	AGGTGCGCGC	3660
	TGCTTATCTG	TGATTGTTGC	TCACCTGAGT	GTGGCTGATT	GTGTACATCC	AGCAGTTACA	3720
	ATTTTAAAA	ATTATACTTT	TACATTTATT	TTATATTTTT	CTCACCCCCA	GTAATTTTCT	3780
	TCCAAAGAAG	TTACATGTA	ATAAGTAGAA	ATCTGTATA	GGAAAAAAGC	ATTAAAAATA	3840
	CTAATATAAC	TGCTTCATTT	GCTGGGAACC	ATTAAAGTA	ATATAAATTA	GCTTTTTTCCA	3900
55	GAAGGATCCT	TTTGTAGCAG	TGTTTATGAA	TGTAACCCCC	AGCAAAATAT	GGCTATATAT	3960
	TAGGGGAGCC	AGTTTGGAGC	AGAGGCCTGA	AGGTCCCTGC	TATGCAGCCG	TGGCCACAGC	4020
	TGCGAGCCCA	AGCACTGTGG	AGCATCCACA	CCTTTGATGG	CAATGCAGAT	TGGTAGCAGG	4080
	TTCCATAGGC	GTACAAAACA	GTATTAAAGC	TCAGTGTTTT	GCATATTGTT	AGCATTTACA	4140
	AATATTTTTG	CTTTAGTATG	AGGAAAGTAA	GGATGGGCAA	AGAAGCGATC	AAAATAGCTA	4200
60	TTGCTACAAC	ATTTTCGAAA	ACAAAGTTGG	GGCTGTATTT	CTTTAAAAAG	ATAAGCCTCT	4260
	AAAAATGCTT	GGCAAAAAAA	ATATAGTGTT	AAAATAGGCC	AGTGATATTA	ATGAGAAAAT	4320
	GAAAGTATGT	ATCAGGAATA	AAGTGATATT	GCATAGGAGT	ATTGTATTTT	TATGAATTTT	4380
	ATGCCAGTTG	TTTACATGTA	CTATATATGT	TAAATTAAAA	AAAATCATGA	GAAATG	

Seq ID NO: 499 Protein sequence  
Protein Accession #: BAA74900.1

	1	11	21	31	41	51	
70	PLVINTLKRF	NLYPEVILAS	WYRIYTKIMD	LIGIQTKICW	TVTRGEGLSP	IESCEGLGDP	60
	ACFYVAVIFI	LNGLMMALFF	IYGYLSGSR	LGGLVTVLCF	FFNHGECTRV	MWTPPLRESF	120
	SYFPLVLQML	LVTHILRATK	LYRGLIALC	ISNVFFMLPW	QFAQFVLLTQ	IASLFAVYVV	180
	GYPIDICKLRK	IIYIHMISLA	LCFVLMFGNS	MLTSSYASS	LVIIWGILAM	KPHFLKINVS	240
	ELSLWVIQGC	FWLFGTVILK	YLTSKIFGIA	DDAHIGNLLT	SKFFSYKDFD	TLTYTCAAEF	300
75	DFMEKETPLR	YTKTLLLPVV	LVVFVAIVRK	IISDMWGVLA	KQQTHVRKHQ	FDHGELVYHA	360
	LQLLAYTALG	ILIMRLKLF	TFHMCVMASL	ICSRQLFGWL	FCKVHPGAI	FAILAAMSIQ	420
	GSANLQTQWN	IVGEFSNLPO	EELIEWIKYS	TKPDAVFAGA	MPTMASVKLS	ALRPIVNHPP	480
	YEDAGLRART	KIVYSMSYRK	AAEEVKRELI	KLKVNYYILE	ESWCVRRSKP	GCSMPFIWDV	540
	EDPANAGKTP	LCNLLVKDSK	PHFTTVFQNS	VYKVLLEVKE			

Seq ID NO: 500 DNA sequence  
Nucleic Acid Accession #: NM\_001275.1  
Coding sequence: 127..1278

	1	11	21	31	41	51	
85	AGTGGAGTGG	GACAGGTATA	TAAAGGAAGT	ACAGGGCCTG	GGGAAGAGGC	CCTGTCTAGG	60
	TAGCTGGCAC	CAGGAGCCGT	GGGCAAGGGA	AGAGGCCACA	CCCTGCCCTG	CTCTGCTGCA	120



5  
10  
15  
20  
25  
30

```

GCCAGAATGG GTGTGAAGGC GTCTCAAACA GGCTTTGTGG TCCTGGTGCT GCTCCAGTGC 180
TGCTCTGCAT ACAAACCTGGT CTGCTACTAC ACCAGCTGGT CCCAGTACCG GGAAGGCGAT 240
GGGAGCTGCT TCCCAGATGC CTTTGACCGC TTCCTCTGTA CCCACATCAT CTACAGCTTT 300
GCCAATATAA GCAACGATCA CATCGACACC TGGGAGTGGG ATGATGTGAC GCTCTACGGC 360
ATGCTCAACA CACTCAAGAA CAGGAACCCC AACCTGAAGA CTCTCTTGTC TGTCCGAGGA 420
TGGAACCTTG GGTCTCAAAG ATTTTCCAAG ATAGCCTCCA ACACCCAGAG TCGCCGGACT 480
TTCATCAAGT CAGTACCGCC ATTCCTGCGC ACCCATGGCT TTGATGGGCT GGACCTTGCC 540
TGGCTCTACC CTGGACGGAG AGACAAACAG CATTTTACCA CCCTAATCAA GGAAATGAAG 600
GCCGAATTTA TAAAGGAAGC CCAGCCAGGG AAAAAGCAGC TCCTGCTCAG CGCAGCACTG 660
TCTGCGGGGA AGGTCACCAT TGACAGCAGC TATGACATTG CCAAGATATC CCAACACCTG 720
GATTTTCATTA GCATCATGAC CTACGATTTT CATGGAGCCT GGCGTGGGAC CACAGGCCAT 780
CACAGTCCCC TGTTCCGAGG TCAGGAGGAT GCAAGTCCTG ACAGATTCAG CAACACTGAC 840
TATGCTGTGG GGTACATGTT GAGGCTGGGG GCTCCTGCCA GTAAGCTGGT GATGGGCATC 900
CCCACCTTCG GGAGGAGCTT CACTCTGGCT TCTTCTGAGA CTGGTGTGGG AGCCCCAATC 960
TCAGGACCGG GAATTCACGG CCGGTTACAC AAGGAGGCAG GGACCCTTGC CTACTATGAG 1020
ATCTGTGACT TCCTCCGCGG AGCCACAGTC CATAGAACCC TCGGCCAGCA GGTCCCCTAT 1080
GCCACCAAGG GCAACCAAGT GGTAGGATAC GACGACCAGG AAAGCGTCAA AAGCAAGGTG 1140
CAGTACCTGA AGGATAGGCA GCTGGCAGGC GCCATGGTAT GGGCCCTGGA CCTGGATGAC 1200
TTCCAGGGCT CCTTCTGCGG CCAGGATCTG CGCTTCCCTC TCACCAATGC CATCAAGGAT 1260
GCACTCGCTG CAACGTAGCC CTCTGTTCTG CACACAGCAC GGGGGCCAAG GATGCCCCGT 1320
CCCCCTCTGG CTCCAGCTGG CCGGGAGCCT GATCACCTGC CCTGCTGAGT CCCAGGCTGA 1380
GCCTCAGTCT CCCTCCCTTG GGGCCTATGC AGAGGTCCAC AACACACAGA TTTGAGCTCA 1440
GCCCTGGTGG GCAGAGAGGT AGGGATGGGG CTGTGGGGAT AGTGAGGCAT CGCAATGTAA 1500
GACTCGGGAT TAGTACACAC TTGTTGATGA TTAATGGAAA TGTTTACAGA TCCCCAAGCC 1560
TGGCAAGGGA ATTTCTTCAA CTCCCTGCCC CCTAGCCCTC CTTATCAAAG GACACCATT 1620
TGGCAAGCTC TATACCAAG GAGCCAAACA TCCTACAAGA CACAGTGACC ATACTAATTA 1680
TACCCCTGCT AAAGCCAGCT TGAACCTTC ACTTAGGAAC GTAATCGTGT CCCCCTATCT 1740
ACTTCCCCTT CCTAATTCCA CAGCTGCTCA ATAAAGTACA AGAGTTTAAC AGTGTGTTGG 1800
CGCTTTGCTT TGGTCTATCT TTGAGCGCCC ACTAGACCCA CTGGACTCAC CTCCCCATC 1860
TCTTCTGGGT TCCTTCCTCT GAGCCTTGGG ACCCCTGAGC TTGCAGAGAT GAAGGCCGCC 1920
ATGTT

```

Seq ID NO: 501 Protein sequence  
Protein Accession #: NP\_001267.1

35  
40  
45

```

1      11      21      31      41      51
|      |      |      |      |      |
MGVKASQTGF VVLVLLQCCS AYKLVCCYYS WSQYREGDGS CFPDALDRFL CTHIIYSFAN 60
ISNDHIDTWE WNDVTLYGML NTLKNRNPNL KTLLSVGGWN FGSQRFASKIA SNTQSRRTFI 120
KSVPPFLRTH FFDGLDLAWL YPGRRDKQHF TTIKEMKAE FIKEAQPGKK QLLLSAALSA 180
GKVTIDSSYD IAKISQHLDF ISIMTYDFHG AWRGTTGHHS PLFRQEDAS PDRFSNTDYA 240
VGYMLRLGAP ASKLVMGIPF FGRSFTLASS ETGVGAPISG PGIPGRFTKE AGTLAYYEIC 300
DFLRGATVHR TLGQQVPYAT KGNQWVGYYD QESVSKSVQY LKDRQLAGAM VWALDLDDFQ 360
GSFCGQDLRF PLTNAIKDAL AAT

```

Seq ID NO: 502 DNA sequence  
Nucleic Acid Accession #: NM\_006474.1  
Coding sequence: 181..669

50  
55  
60  
65

```

1      11      21      31      41      51
|      |      |      |      |      |
GCTGCCTAGG GTCTGGAAG CTCGGGCACC CTCCCTCTCC GGGGCTCCTG CTCCACCCC 60
TCCGGCCCCC CCACCGTCGC GTCCTCTCAG GCTGGGCCTG TGGCCGCGGT GCTTTTAATT 120
TTCCCCCAGC TCAGATCTT GCTGCTCGGC CCCAGGAGA GCAACAATC AACGGGAACG 180
ATGTGGAAGG TGTCAGCTCT GCTCTTCGTT TTGGGAAGCG CGTCGCTCTG GGTCTTGCA 240
GAAGGAGCCA GCACAGGCCA GCCAGAAGAT GACACTGAGA CTACAGGTTT GGAAGGCGGC 300
GTTGCCATGC CAGGTGCCGA AGATGATGTG GTGACTCCAG GAACCAGCGA AGACCGCTAT 360
AAGTCTGGCT TGACAACTCT GGTGGCAACA AGTGTCAACA GTGTAAACAGG CATTGCGATC 420
GAGGATCTGC CAACCTCAGA AAGCACAGTC CACGCGCAAG AACAAAGTCC AAGCGCCACA 480
GCCTCAAACG TGGCCACCAG TCACTCCACG GAGAAAGTGG ATGGAGACAC ACAGACAACA 540
GTTGAGAAAG ATGGTTTGTG AACAGTGACC CTGGTTGGAA TCATAGTTGG GGTCTTACTA 600
GCCATCGGTT TCATTGGTGG AATCATCGTT GTGGTTATGC GAAAAATGTC GGAAGGTAC 660
TCGCCCTAAA GAGCTGAAGG GTTACGCCCT GCTTGCCAAC GTGCTTTAAA AAAAGACCGT 720
TTCTGACTCT GTGGCCCTGT CCCTGAGCTC GTGGGGAGAA GATGACCCTG GGAACATTTG 780
CGGGCCCATT CAGATTCCAC GGTGACTTTC CGTTTGCCAA ATTAACCGAG GAAAGACCTT 840
TCACCAGATT TGGTTCTTAA ACTTT

```

Seq ID NO: 503 Protein sequence  
Protein Accession #: NP\_006465.1

70  
75

```

1      11      21      31      41      51
|      |      |      |      |      |
MWKVSALLFV LGSASLWVLA EGASTGQPED DTETTGLEGG VAMPGAEDDV VTPGTSEDY 60
KSLGTLTVAT SVNSVTGIRI EDLPTSESTV HAQEQSPSAT ASNVATSHST EKVDGDTQTT 120
VEKDGLSTVT LVGIIVGVLL AIGFIGGIIV VVMRKMSGRY SP

```

Seq ID NO: 504 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 62..895

80  
85

```

1      11      21      31      41      51
|      |      |      |      |      |
CACTGCTCTG AGAATTTGTG AGCAGCCCCT AACAGGCTGT TACTTCACTA CAACTGACGA 60
TATGATCATC TTAATTTACT TATTTCTCTT GCTATGGGAA GACACTCAAG GATGGGGATT 120
CAAGGATGGA ATTTTTCATA ACTCCATATG GCTTGAACGA GCAGCCGGTG TGTACCACAG 180
AGAAGCACGG TCTGGCAAAT ACAAGCTCAC CTACGCAGAA GCTAAGGCGG TGTGTGAATT 240
TGAAGGCGGC CATCTCGCAA CTTACAAGCA GCTAGAGGCA GCCAGAAAAA TTGGATTTC 300

```



TGTCTGTGCT GCTGGATGGA TGGCTAAGGG CAGAGTTGGA TACCCCATG TGAAGCCAGG 360  
 GCCCAACTGT GGATTTGGAA AAAGTGGCAT TATTGATTAT GGAATCCGTC TCAATAGGAG 420  
 TGAAAGATGG GATGCCTATT GCTACAACCC ACACGCAAG GAGTGTGGTG GCGTCTTTAC 480  
 AGATCCAAAG CAAATTTTAA AATCTCCAGG CTTCCCAAAT GAGTACGAAG ATAACCAAAT 540  
 CTGCTACTGG CACATTAGAC TCAAGTATGG TCAGCGTATT CACCTGAGTT TTTTAGATTT 600  
 TGACCTTGAA GATGACCCAG GTTGCTTGGC TGATTATGTT GAAATATATG ACAGTTACGA 660  
 TGATGTCCAT GGCTTTGTGG GAAGATACTG TGGAGATGAG CTTCCAGATG ACATCATCAG 720  
 TACAGGAAAT GTCATGACCT TGAAGTTTCT AAGTGTGCT TCAGTGACAG CTGGAGGTTT 780  
 CCAAATCAAA TATGTTGCAA TGGATCCTGT ATCCAAATCC AGTCAAGGAA AAAATACAAG 840  
 TACTACTTCT ACTGGAAATA AAAACTTTTT AGCTGGAAGA TTTAGCCACT TATAAAAAAA 900  
 AAAAAAAGGA TGATCAAAAC ACACAGTGTT TATGTTGGAA TCTTTTGGAA CTCCTTTGAT 960  
 CTCCTGTGTA TTATTAACAT TTATTTATTA TTTTCTAAA TGTGAAAGCA ATACATAATT 1020  
 TAGGGAAAT TGGAAATAT AGGAACTTT AAACGAGAAA ATGAAACCTC TCATAATCCC 1080  
 ACTGCATAGA AATAACAAGC GTTAACATTT TCATATTTTT TTCTTTCAGT CATTTTTCTA 1140  
 TTTGTGGTAT ATGTATATAT GTACCTATAT GTATTTGCAT TTGAAATTTT GGAATCCTGC 1200  
 TCTATGTACA GTTTTGTATT ATACTTTTTA AATCTTGAAC TTTATAAACA TTTTCTGAAA 1260  
 TCATTGATTA TTCTACAAA ACATGATTTT AAACAGCTGT AAAATATTCT ATGATATGAA 1320  
 TGTTTTATGC ATTATTTAAG CCTGTCTCTA TTGTTGGAAT TTCAGGTCAT TTTTATAAAT 1380  
 ATTGTTGCAA TAAATATCCT TGAACACACA AAAAAAATAA AA

Seq ID NO: 505 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | |  
 MIILIIYLFLL LWEDTQGWGF KDGIHFHNSIW LERAAGVYHR EARSQKYKLT YAEAKAVCEF 60  
 EGGHLATYKQ LEARKIGFH VCAAGWMAKG RVGYPIVKPG PNCQFGKGTGI IDYGIRLNRS 120  
 ERWDAYCYNP HAKECGGVFT DPKQIFKSPG FPNEDNQI CYWHIRLKYG QRIHLSFLDF 180  
 DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELPDDIIS TGNVMTLKFLL SDASVTAGGF 240  
 QIKYVAMDPV SKSSQGNKTS TTSTGNKNFL AGRFSLH

Seq ID NO: 506 DNA sequence  
 Nucleic Acid Accession #: NM\_007115.1  
 Coding sequence: 69..902

1 11 21 31 41 51  
 | | | | |  
 GAATTCGCAC TGCTCTGAGA ATTTGTGAGC AGCCCCTAAC AGGCTGTTAC TTCACTACAA 60  
 CTGACGATAT GATCATCTTA ATTTACTTAT TTCTCTTGCT ATGGGAAGAC ACTCAAGGAT 120  
 GGGGATTCAA GGATGGAATT TTTCATAACT CCATATGGCT TGAACGAGCA GCCGGTGTGT 180  
 ACCACAGAGA AGCAGGTCTT GGCATAATA AGCTCACCTA CGCAGAAGCT AAGGCGGTGT 240  
 GTGAATTTGA AGGCGGCCAT CTCGCAACTT ACAAGCAGCT AGAGGCAGCC AGAAAAATTG 300  
 GATTTTCATGT CTGTGCTGCT GGATGGATGG CTAAGGGCAG AGTTGGATAC CCCATTGTGA 360  
 AGCCAGGGCC CAAGTATGTA TTTGGAAAAA CTGGCATTAT TGATTATGGA ATCCGTCTCA 420  
 ATAGGAGTGA AAGATGGGAT GCCTATTGCT ACAACCCACA CGCAAAGGAG TGTGGTGGCG 480  
 TCTTTACAGA TCCAAAGCGA ATTTTAAAT CTCCAGGCTT CCCAAATGAG TACGAAGATA 540  
 ACCAAATCTG CTACTGGCAC ATTAGACTCA AGTATGGTCA GCGTATTCAC CTGAGTTTTT 600  
 TAGATTTTGA CCTTGAAGAT GACCCAGGTT GCTTGGCTGA TTATGTTGAA ATATATGACA 660  
 GTTACGATGA TGTCCATGGC TTTGTGGGAA GATACTGTGG AGATGAGCTT CCAGATGACA 720  
 TCATCAGTAC AGGAAATGTC ATGACCTTGA AGTTTCTAAG TGATGCTTCA GTGACAGCTG 780  
 GAGGTTTCCA AATCAAATAT GTTGCAATGG ATCCTGTATC CAAATCCAGT CAAGGAAAAA 840  
 ATACAAGTAC TACTTCTACT GGAAATACAA ACTTTTTAGC TGGAAGATTT AGCCACTTAT 900  
 AAAAAAATAA AAAGATGATC AAAACACACA GTGTTTATGT TGGAAATCTT TGGAACTCCT 960  
 TTGATCTCAC TGTTATTATT AACATTTATT TATTATTTTT CTAAATGTGA AAGAAATACA 1020  
 TAATTTAGGG AAAATTGGAA AATATAGGAA ACTTTAAACG AGAAAAATGAA ACCTCTCATA 1080  
 ATCCCACTGC ATAGAAATAA CAAGCGTTAA CATTTTCATA TTTTCTCTT TCAGTCATTT 1140  
 TTGTATTTGT GGTATATGTA TATATGTACC TATATGTATT TGCATTTGAA ATTTTGGAA 1200  
 CCTGCTCTAT GTACAGTTT GTATTATCT TTTTAAATCT TGAACCTTAT GAACATTTTC 1260  
 TGAAATCATT GATTATCTA CAAAACATG ATTTTAAACA GCTGTAAAAT ATTCTATGAT 1320  
 ATGAATGTTT TATGCATTAT TTAAGCCTGT CTCTATTGTT GGAATTTTCA GTCATTTTCA 1380  
 TAAATATTGT TGCAATAAAT ATCCTTCGGA ATTC

Seq ID NO: 507 Protein sequence  
 Protein Accession #: NP\_009046.1

1 11 21 31 41 51  
 | | | | |  
 MIILIIYLFLL LWEDTQGWGF KDGIHFHNSIW LERAAGVYHR EARSQKYKLT YAEAKAVCEF 60  
 EGGHLATYKQ LEARKIGFH VCAAGWMAKG RVGYPIVKPG PNXKFGKGTGI IDYGIRLNRS 120  
 ERWDAYCYNP HAKECGGVFT DPKRIFKSPG FPNEDNQI CYWHIRLKYG QRIHLSFLDF 180  
 DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELPDDIIS TGNVMTLKFLL SDASVTAGGF 240  
 QIKYVAMDPV SKSSQGNKTS TTSTGNKNFL AGRFSLH

Seq ID NO: 508 DNA sequence  
 Nucleic Acid Accession #: NM\_001044.1  
 Coding sequence: 129..1991

1 11 21 31 41 51  
 | | | | |  
 ACCGCTCCGG AGCGGGAGGG GAGGCTTCGC GGAACGCTCT CGGCGCCAGG ACTCGCGTGC 60  
 AAAGCCCAGG CCCGGGCGGC CAGACCAAGA GGGAAGAAGC ACAGAATTCC TCAACTCCCA 120  
 GTGTGCCCAT GAGTAAGAGC AAATGCTCCG TGGGACTCAT GTCTTCCGTG GTGGCCCCGG 180  
 CTAAGGAGCC CAATGCCGTG GCGCCGAGAG AGGTGAGGCT CATCTTGTG AAGGAGCAGA 240  
 ACGGAGTGCA GCTCACCAGC TCCACCCTCA CCAACCCGCG GCAGAGCCCC GTGGAGGCC 300  
 AGGATCGGGA GACCTGGGGC AAGAAGATCG ACTTTCTCCT GTCCGTCAAT GGCTTTGCTG 360  
 TGGACCTGGC CAACGTCTGG CGGTTCCCCT ACCTGTGCTA CAAAATGGT GCGGTGCCT 420



5	TCCTGGTCCC	CTACCTGCTC	TTCATGGTCA	TTGCTGGGAT	GCCACTTTTC	TACATGGAGC	480
	TGGCCCTCGG	CCAGTTCAAC	AGGGAAGGGG	CCGCTGGTGT	CTGGAAGATC	TGCCCCATAC	540
	TGAAAGGTGT	GGGCTTCACG	GTCATCCTCA	TCTCACTGTA	TGTCGGCTTC	TTCTACAACG	600
	TCATCATCGC	CTGGGCGCTG	CACATATCTCT	TCTCCTCCTT	CACCACGGAG	CTCCCTGGA	660
	TCCACTGCAA	CAACTCCTGG	AACAGCCCCA	ACTGCTCGGA	TGCCCCATCCT	GGTGACTCCA	720
	GTGGAGACAG	CTCGGGCCTC	AACGACACTT	TTGGGACCAC	ACCTGCTGCC	GAGTACTTTG	780
	AACGTGGCGT	GCTGCACCTC	CACCAGAGCC	ATGGCATCGA	CGACCTGGGG	CCTCCGCGGT	840
	GGCAGCTCAC	AGCCTGCCTG	GTGCTGGTCA	TCTGTCTGCT	CTACTTCAGC	CTCTGGAAGG	900
10	GCGTGAAGAC	CTCAGGGAAG	GTGGTATGGA	TCACAGCCAC	CATGCCATAC	GTGGTCCTCA	960
	CTGCCCTGCT	CCTGCGTGGG	GTCACCCCTC	CTGGAGCCAT	AGACGGCATC	AGAGCATACC	1020
	TGAGCGTTGA	CTTCTACCGG	CTCTGCGAGG	CGTCTGTTTG	GATTGACGCG	GCCACCCAGG	1080
	TGTGCTTCTC	CCTGGGCGTG	GGGTTCCGGG	TGCTGATCGC	CTTCTCCAGC	TACAACAAGT	1140
	TCACCAACAA	CTGTACAGG	GACGCGATTG	TCACCACCTC	CATCAACTCC	CTGACGAGCT	1200
	TCTCCTCCGG	CTTCGTGCTC	TTCTCCTTCC	TGGGTACAT	GGCACAGAAG	CACAGTGTGC	1260
15	CCATCGGGGA	CGTGGCCAAG	GACGGGCCAG	GGCTGATCTT	CATCATCTAC	CCGGAAGCCA	1320
	TCGCCACGCT	CCCTCTGTCC	TCAGCCTGGG	CCGTGGTCTT	CTTCATCATG	CTGCTCACCC	1380
	TGGGTATCGA	CAGCGCCATG	GGTGGTATGG	AGTCAGTGAT	CACCGGGCTC	ATCGATGAGT	1440
	TCCAGCTGCT	GCACAGACAC	CGTGAGCTCT	TCACGCTCTT	CATCGTCTTG	GCGACCTTCC	1500
20	TCCTGTCCCT	GTTCTGCGTC	ACCAACGGTG	GCATCTACGT	CTTCACGCTC	CTGGACCATT	1560
	TTGCAGCCGG	CACGTCCATC	CTCTTTGGAG	TGCTCATCGA	AGCCATCGGA	GTGGCCTGGT	1620
	TCTATGGTGT	TGGGCAGTTC	AGCGACGACA	TCCAGCAGAT	GACCGGCGAG	CGGCCAGGCC	1680
	TGTACTGGCG	GCTGTGCTGG	AAGCTGGTCA	GCCCCTGCTT	TCTCCTGTTC	GTGGTCTGTG	1740
	TCAGCATTGT	GACCTTCAGA	CCCCCCCCACT	ACGGAGCCTA	CATCTTCCCC	GACTGGGCCA	1800
25	ACGCGCTGGG	CTGGGTCAATC	GCCACATCCT	CCATGGCCAT	GGTGCCCATC	TATGCGGCCT	1860
	ACAAGTTCTG	CAGCCTGCCT	GGGTCTTTTC	GAGAGAAACT	GGCCTACGCC	ATTGCACCCG	1920
	AGAAGGACCG	TGAGCTGGTG	GACAGAGGGG	AGGTGCGCCA	GTTACGCTC	CGCCACTGGC	1980
	TCAAGGTGTA	GAGGAGCAG	AGACGAAGC	CCAGGAAGT	CATCCTGCAA	TGGGAGAGAC	2040
	ACGAACAAC	CAAGGAAATC	TAAAGTTTCGA	GAGAAAGGAG	GGCAACTTCT	ACTCTTCAAC	2100
30	CTCTACTGAA	AACACAAACA	ACAAAGCAGA	AGACTCCTCT	CTTCTGACTG	TTTACACCTT	2160
	TCCGTGCCGG	GAGCGCACCT	CGCCGTGTCT	TGTGTTGCTG	TAATAACGAC	GATAGTCTGT	2220
	GCAGCGAGGT	CCACCCCGTT	GTTGTCCCTG	CAGGGCAGAA	AAACGTCTAA	CTTCATGCTG	2280
	TCTGTGTGAG	GCTCCCTCCC	TCCCTGCTCC	CTGCTCCCGG	CTCTGAGGCT	GCCCCAGGGG	2340
	CACGTGTGTT	TCAGCGGGG	ATCAGCATCC	TTGTAGACGC	ACCTGCTGAG	AATCCCCGTG	2400
35	CTCAGTAGAG	CTTCTTAGAC	CATTACTTTT	GCCCATATTA	AAAAGCCAAAG	TGTCCTGCTT	2460
	GGTTTAGCTG	TGCAGAAAGT	GAAATGGAGG	AAACCACAAA	TTCATGCAAA	GTCCTTTCCC	2520
	GATGCGTGGC	TCCCAGCAGA	GGCCGTAAAT	TGAGCGTTCA	GTTGACACAT	TGCACACACA	2580
	GTCTGTTTCA	AGGCATTGGA	GGATGGGGGT	CCTGGTATGT	CTCACCAGGA	AATTCTGTTT	2640
	ATGTTCTTGC	AGCAGAGAGA	AATAAACTC	CTTGAAACCA	GCTCAGGCTA	CTGCCACTCA	2700
40	GGCAGCCTGT	GGGTCCTTGT	GGTGTAGGGA	ACCGCCTGAG	AGGAGCGTGT	CCTATCCCCG	2760
	GACGCATGCA	GGGCCCCAC	AGGAGCGTGT	CCTATCCCCG	GACGCATGCA	GGGCCCCAC	2820
	AGGAGCATGT	CCTATCCCTG	GACGCATGCA	GGGCCCCAC	AGGAGCGTGT	ACTACCCAG	2880
	AACGCATGCA	GGGCCCCAC	AGGAGCGTGT	ACTACCCAG	GACGCATGCA	GGGCCCCAC	2940
	TGGAGCGTGT	ACTACCCAG	GACGCATGCA	GGGCCCCAC	AGGAGCGTGT	CCTATCCCCG	3000
45	GACCGGACGC	ATGCAGGGCC	CCCACAGGAG	CGTGTAATAC	CCCAGGACGC	ATGCAGGGCC	3060
	CCCACAGGAG	CGTGTAATAC	CCCAGGATGC	ATGCAGGGCC	CCCACAGGAG	CGTGTAATAC	3120
	CCCAGGACGC	ATGCAGGGCC	CCCAGGATGC	CAGCCTGCAG	ACCAACACTC	TGCCTGGCCT	3180
	TGAGCCGTGA	CCTCCAGGAA	GGGACCCAC	TGGAATTTTA	TTTCTCTCAG	GTGCGTGCCA	3240
	CATCAATAAC	AACAGTTTTT	ATGTTTGCGA	ATGGCTTTTT	AAAATCATAT	TTACCTGTGA	3300
50	ATCAAAACAA	ATTCAAGAAT	GCAGTATCCG	CGAGCCTGCT	TGCTGATATT	GCAGTTTTTG	3360
	TTTACAAGAA	TAATTAGCAA	TACTGAGTGA	AGGATGTTGG	CCAAAAGCTG	CTTTCCATGG	3420
	CACACTGCCC	TCTGCCACTG	ACAGGAAAGT	GGATGCCATA	GTTTGAATTC	ATGCCTCAAG	3480
	TCCGTGGGCC	TGCTTACGTG	CTGCCCGAGG	GCAGGGGCCG	TGCAGGGCCA	GTGATGGCTG	3540
	TCCCTGCAA	GTGGACGTGG	GCTCCAGGGA	CTGGAGTGTA	ATGCTCGGTG	GGAGCCGTCA	3600
55	GCCTGTGAAC	TGCCAGGCAG	CTGCAGTTAG	CACAGAGGAT	GGCTTCCCCA	TTGCCTTCTG	3660
	GGGAGGGACA	CAGAGGACGG	CTTCCCCATC	GCCTTCTGGC	CGCTGCAGTC	AGCACAGAGA	3720
	GCGGCTTCCC	CATTGCCTTC	TGGGAGGGGA	CACAGAGGAC	AGTTTCCCCA	TGCTCTTCTG	3780
	GTTGTTGAAG	ACAGCACAGA	GAGCGGCTTC	CCCATCGCCT	TCTGGGGAGG	GGCTCCGTGT	3840
60	AGCAACCCAG	GTGTTGTCCG	TGTCTGTTGA	CCAATCTCTA	TTGAGCATCG	TGTGGGTCCC	3900
	TAAGCACAAAT	AAAAGACATC	CACAATGGAA	AAAAAAAAG	GAATTC		

Seq ID NO: 509 Protein sequence  
Protein Accession #: NP\_001035.1

65	1	11	21	31	41	51	
	MSKSKCSVGL	MSSVAPAKE	PNAVGPKEVE	LILVKEQNGV	QLTSSTLTNP	RQSPVEAQDR	60
	ETWGGKIDFL	LSVIGFAVDL	ANVWRFPYLC	YKNGGGAFLV	PYLLEFMVIAG	MPLFYMEALAL	120
	GQFNREGAAG	VWKICPILKG	VGFTVILISL	YVGFFYNVII	AWALHYLFSS	FTTELPWIHC	180
70	NNSWNSPNC	DAHPSDSSGD	SSGLNDTFGT	TPAAEYFERG	VLHLHQSHGI	DDLGPWRWL	240
	TACLVLVIVL	LYFSLWKGVK	TSKGVVWITA	TMPIVVLTLA	LLRGVTLPGA	IDGIRAYLSV	300
	DFYRLCEASV	WIDAAQVCF	SLGVGFVGLI	AFSSYNKFTN	NCYRDAIVTT	SINSLTSFSS	360
	GFVVFSFLGY	MAQKHSVPIG	DVAKDGPGLI	FIIYPEAIAT	LPLSSAWAVV	FFIMLLTLGI	420
	DSAMGMESV	ITGLIDEFQL	LHRHRELFTL	FIVLATFLLS	LFCVTNGGIY	VFTLLDHFAA	480
75	GTSILFGVLI	EAIGVAWFY	VGQFSDDIQ	MTGQRPSLYW	RLCWKLVSFC	FLLFVVVSI	540
	VTERPPHYGA	YIFPDWANAL	GWVIATSSMA	MVPIYAAYKF	CSLPGSFREK	LAYAIAPKED	600
	RELVDREVR	QFTLRHHLKV					

Seq ID NO: 510 DNA sequence  
Nucleic Acid Accession #: NM\_001216.1  
Coding sequence: 43..1422

85	1	11	21	31	41	51	
	GCCCGTACAC	ACCGTGTGCT	GGGACACCCC	ACAGTCAGCC	GCATGGCTCC	CCTGTGCCCC	60
	AGCCCCCTGGC	TCCCTCTGTT	GATCCCGGCC	CCTGCTCCAG	GCCTCACTGT	GCAACTGCTG	120
	CTGTCACTGC	TGCTTCTGAT	GCCTGTCCAT	CCCCAGAGGT	TGCCCCGGAT	GCAGGAGGAT	180
	TCCCCCTTGG	GAGGAGGCTC	TTCTGGGGAA	GATGACCCAC	TGGGCGAGGA	GGATCTGCCC	240



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

AGTGAAGAGG ATTCACCCAG AGAGGAGGAT CCACCCGGAG AGGAGGATCT ACCTGGAGAG 300  
GAGGATCTAC CTGGAGAGGA GGATCTACCT GAAGTTAAGC CTAAATCAGA AGAAGAGGGC 360  
TCCCTGAAGT TAGAGGATCT ACCTACTGTT GAGGCTCCTG GAGATCCTCA AGAACCCAG 420  
AATAATGCCC ACAGGGACAA AGAAGGGGAT GACCAGAGTC ATTGGCGCTA TGGAGGCGAC 480  
CCGCCCTGGC CCCGGGTGTC CCCAGCCTGC GCGGGCCGCT TCCAGTCCCC GGTGGATATC 540  
CGCCCCCAGC TCGCCGCCTT CTGCCCGGCC CTGCGCCCCC TGGAACTCCT GGGCTTCCAG 600  
CTCCCGCCGC TCCCAGAACT GCGCCTGCGC AACAATGGCC ACAGTGTGCA ACTGACCCTG 660  
CCTCCTGGGC TAGAGATGGC TCTGGGTCCC GGGCGGAGT ACCGGGCTCT GCAGCTGCAT 720  
CTGCACTGGG GGGCTGCAGG TCGTCCGGGC TCGGAGCACA CTGTGGAAGG CCACCGTTTC 780  
CCTGCCGAGA TCCACGTGGT TCACCTCAGC ACCGCCTTTG CCAGAGTTGA CGAGGCCTTG 840  
GGGCGCCCGG GAGGCCTGGC CGTGTGGGCC GCCTTTCTGG AGGAGGGCCC GGAAGAAAAC 900  
AGTGCCTATG AGCAGTTGCT GTCTCGCTTG GAAGAAATCG CTGAGGAAGG CTCAGAGACT 960  
CAGGTCCAG GACTGGACAT ATCTGCACTC CTGCCCTCTG ACTTCAGCCG CTACTTCCAA 1020  
TATGAGGGGT CTCTGACTAC ACCGCCCTGT GCCCAGGGTG TCATCTGGAC TGTGTTTAAAC 1080  
CAGACAGTGA TGCTGAGTGC TAAGCAGCTC CACACCCTCT CTGACACCCT GTGGGGACCT 1140  
GGTGA CTCTC GGCTACAGCT GAACTTCCGA GCGACGCAGC CTTTGAATGG GCGAGTGATT 1200  
GAGGCCTCCT TCCCTGCTGG AGTGGACAGC AGTCCTCGGG CTGCTGAGCC AGTCCAGCTG 1260  
AATTCTGCTG TGGCTGCTGG TGACATCCTA GCCCTGGTTT TTGGCCTCCT TTTTGTCTGTC 1320  
ACCAGCGTCG CGTTCCTTGT GCAGATGAGA AGGCAGCACA GAAGGGGAAC CAAAGGGGGT 1380  
GTGAGCTACC GCCCAGCAGA GGTAGCCGCT ACTGGAGCCT AGAGGCTGGA TCTTGGAGAA 1440  
TGTGAGAAGC CAGCCAGAGG CATCTGAGGG GGAGCCGTA ACTGTCCTGT CCTGCTCATT 1500  
ATGCCACTTC CTTTAACTG CCAAGAAATT TTTTAAATA AATATTTATA AT

Seq ID NO: 511 Protein sequence  
Protein Accession #: NP\_001207.1

1 11 21 31 41 51  
MAPLCPSPWL PLLIPAPAPG LTVQLLLSLL LLMPVHPQRL PRMQEDSPLG GGSSGEDDPL 60  
GEEDLPSEED SPREEDPPGE EDLPGEEDLP GEEDLPPEVKP KSEEEGSLKL EDLPTVEAPG 120  
DPQEPQNNAH RDKEGDDQSH WRYGGDPPWP RVSPACAGRF QSPVDIRPQL AAFCPALRPL 180  
ELLGFQLPPL PELRLRNNGH SVQLTLPPLG EMALGPGREY RALQLHLHWG AAGRPGSEHT 240  
VEGHRFPAEI HVVHLSTAFV RVDEALGRPG GLAVLAAFL EGPENSAYE QLLSRLEBIA 300  
EEGSETQVPG LDISALLPSD FSRYFQYEGS LTTFPCAQGV IWTVFNQTMV LSAKQLHTLS 360  
DTLWGPDSR LQLNFRATQP LNGRVIEASF PAGVDSSPRA AEPVQLNSCL AAGDILALVF 420  
GLLFAVTSVA FLVQMRQRH RGTGKGVSYR PAEVAETGA

Seq ID NO: 512 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..3978

45  
50  
55  
60  
65  
70  
75  
80  
85

1 11 21 31 41 51  
ATGGTGGGTG AAGGACCCTA CCTTATCTCA GATCTGGACC AGCGAGGCCG GCGGAGATCC 60  
TTTGACAGAA GATATGACCC CAGCCTGAGC ACCATGATCC CAGTGCAGCC CTGTGCAAGG 120  
TTAGCACCCA ACCCGGTGGA TGATGCGGG CTACTCTCCT TCGCCACATT TTCTGGCTC 180  
ACGCCGGTGA TGGTGAAAGG CTACCGGCAA AGGCTGACCG TAGACACCCT GCGCCCATTC 240  
TCGACATATG ACTCATCTGA CACCAATGCC AAAAGATTTT GAGTCCTTTG GGATGAAGAG 300  
GTAGCAAGGG TGGGTCTCTG GAAGGCCTCT CTGAGCCACG TGGTGTGGAA ATCCAGAGG 360  
ACACGCGTGT TGATGGACAT CGTGGCCAA C ATCCTGTGCA TCATCATGGC AGCCATAGGG 420  
CCGACAGTTC TCATTACCA AATCCTCCAG CAGACTGAGA GGACCTCTGG GAAAGTCTGG 480  
GTTGGCATTG GACTGTGCAT AGCCCTTTT GGCACCGAGT TTACCAAAGT CTCTTTTGG 540  
GCCCTTGCCCT GGGCCATCAA CTACCGCACG GCCATCCGGT TGAAGGTGGC GCTCTCCACC 600  
TTGGTTTTTG AAAACCTAGT GTCCTTCAAG ACATTGACCC ACATCTCTGT TGGCGAGGTG 660  
CTCAATATAC TGTCAAGTGA TAGCTATTCT TTGTTTGAAG CTGCCTTGTT TTGTCTTTG 720  
CCAGCCACCA TCCCGATCCT AATGGTCTTT TGTGCGGCGT ACGCCTTTTT CATTCCTGGG 780  
CCCACAGCTC TCATCGGGAT ATCAGTGTAT GTCATATTCA TACCCGTCCA GATGTTTATG 840  
GCCAAGCTCA ATTCAGCTTT CCGAAGGTCA GCAATTTTGG TGACAGACAA GCGAGTTCAG 900  
ACAAATGAATG AGTTTCTGAC CTGCATCAGG CTGACAAAA TGTATGCCTG GGAGAAATCT 960  
TTTACCAACA CTATCCAAGA TATAAGAAGG AGGGAAAGAA AATTACTGGA AAAAGCTGGA 1020  
TTTGTCCAAA GTGGAAACTC TGCCCTGGCC CCCATCGTGT CCACCATAGC CATCGTCTG 1080  
ACATTATCCT GCCACATCCT CCTGAGACGC AAACCTACCG CACCCGTGGC ATTTAGTGTG 1140  
ATTGCCATGT TTAATGTAAT GAAGTTTTTC ATTGCAATCT TGCCCTTCTC CATCAAAGCA 1200  
ATGGCTGAAG CGAATGTCTC TCTAAGGAGA ATGAAGAAAA TTCTCATAGA TAAAAGCCCC 1260  
CCATCTTACA TCACCAACC AGAAGACCCA GATACTGTCT TGCTTTTAGC AAATGCCACC 1320  
TTGACATGGG AGCATGAAGC CAGCAGGAAA AGTACCCCAA AGAAATTGCA GAACCAGAAA 1380  
AGGCATTTAT GCAAGAAACA GAGGTGAGAG GCATACAGTG AGAGGAGTCC ACCAGCCAAG 1440  
GGAGCCACTG GCCCAGAGGA GCAAAGTGAC AGCCTCAAAT CGGTTCTGCA CAGCATAAGC 1500  
TTTGTGGTGA GAAAGTTATG TCGTTATCCC GAAGCCCAGC TCCTGGCTTG GAGGTGGCCA 1560  
GCAGTGTGTTG TTGGGAGAAT CATCAGAGGA TACAGGCCTC ATGGATTTTC TGCTAAAGAC 1620  
AAGGATGAAT CTAGAAGGCT TCTTACTTGG CCCCAGGAAG TGGATAGGAC TCAAAGGGCA 1680  
GCCAAATACC TGGGGAAGAT CTTGGGAATA TGTGGGAATG TGGGAAGTGG AAAGAGCTCC 1740  
CTCCTTGACG CTCTCCTAGG ACAGATGCAG CTGCAGAAA GGGTGGTGGC AGTCAATGGA 1800  
ACTTTGGCCT ACGTTTCACA GCAGGCATGG ATCTTTCATG GAAATGTGAG AGAAAAACATA 1860  
CTCTTTGGAG AAAAGTATGA TCACCAAAGG TATCAGCACA CAGTCCGCGT CTGTGGCCTC 1920  
CAGAAGGACC TGAGCAACCT CCCCTATGGA GACCTGACTG AGATTGGGGA GCGGGGCCCT 1980  
AACCTCTCTG GGGGGCAGAG GCAGAGGATT AGCCTGGGCC GCGCTGTCTA CTCCGACCGT 2040  
CAGCTCTACC TGCTGGACGA CCCCCTGTG GCGGTGACG CCCACGTGGG GAAGCACGTC 2100  
TTTGAGGAGT GCATTAAGAA GACGCTCAGG GGAAAGACAG TCGTCTGGT GACCCACCAG 2160  
CTACAGTTCT TAGAGTCTTG TGATGAAGTT ATTTTATTAG AAGATGGAGA GATTTGTGAA 2220  
AAGGGAACCC ACAAGGAGTT AATGGAGGAG AGAGGGCGCT ATGCAAAACT GATTCACAAC 2280  
CTGCGAGGAT TGCAGTTCAA GGATCCTGAA CACCTTTACA ATGCAGCAAT GGTGGAAGCC 2340  
TTCAAGGAGA GCCCTGCTGA GAGAGAGGAA GATGCTGGTA TAATCGGGTA CCTCCTTTCT 2400  
CTCTTCACTG GTTCTCTCTT CCTCCTGATG ATTGGCAGCG CTGCCTTCAG CAACTGGTGG 2460  
CTGGGTCTCT GGTGGACAA GGGCTCACGG ATGACCTGTG GGCCCCAGGG CAACAGGACC 2520  
ATGTGTGAGG TCGGCGCGGT GCTGGCAGAC ATCGGTGAGC ATGTGTACCA GTGGGTGTAC 2580  
ACTGCAAGCA TGGTGTTCAT GCTGGTGTGTT GCGGTCACCA AAGGCTTCGT CTTACCAAG 2640



ACCACACTGA TGGCATCCTC CTCTCTGCAT GACACGGTGT TTGATAAGAT CTAAAGAGC 2700  
 CCAATGAGTT TCTTTGACAC GACTCCCACT GGCAGGCTAA TGAACCGTTT TTCCAAGGAT 2760  
 ATGGACGAGC TGGATGTGAG GCTGCCGTTT CACGCAGAGA ACTTTCTGCA GCAGTTTTTT 2820  
 ATGGTGGTGT TTATTCTCGT GATCTTGGCT GCTGTGTTTC CTGCTGTCTT TTTAGTCGTG 2880  
 5 GCCAGCCTTG CTGTAGGCTT CTTTATTCTG TTACGCATTT TCCACAGAGG AGTCCAGGAG 2940  
 CTCAAGAAGG TGGAGAATGT CAGCCGGTCA CCCTGGTTCA CCCACATCAC CTCCTCCATG 3000  
 CAGGGCCTGG GCATCATTCA CGCCTATGGC AAGAAGGAGA GCTGCATCAC CTATACTTCA 3060  
 TCCAAAGGCC TGTCATTGTC ATACATCATC CAGCTGAGCG GACTGCTCCA AGTGTGTGTG 3120  
 10 CGAACGGGAA CAGAGACGCA AGCCAAATTC ACCTCCGTGG AGCTGCTCAG GGAATACATT 3180  
 TCGACCTGTG TTCCTGAATG CACTCATCCC CTCAAAGTGG GGACCTGTCC CAAGGACTGG 3240  
 CCCAGCTGTG GGGAGATCAC CTTTCAAGAG TATCAGATGA GATACAGAGA CAACACCCCC 3300  
 CTTGTTCTCG ACAGCCTGAA CTTGAACATA CAAAGTGGGC AGACAGTCGG GATTGTTGGA 3360  
 AGAACAGGTT CCGGAAAGTC ATCGTTAGGA ATGGCTTTGT TTCGTCTGGT GGAGCCAGCC 3420  
 AGTGGCACAA TCTTTATTGA TGAGGTGGAT ATCTGCATTG TCAGCTTGGG AGACCTCAGA 3480  
 15 ACCAAGCTGA CTGTGATCCC ACAGGATCCT GTCTGTGTTG TAGGTACAGT AAGGTACAAC 3540  
 TTGGATCCCT TTGAGAGTCA CACCGATGAG ATGCTCTGGC AGGTCTCTGA GAGAACATTC 3600  
 ATGAGAGACA CAATAATGAA ACTCCCAGAA AAATTACAGG CAGAAGTCAC AGAAAATGGA 3660  
 GAAAACCTCT CAGTAGGGGA ACGTCAGCTG CTTTGTGTGG CCCGAGCTCT TCTCCGTAAT 3720  
 TCAAAGATCA TTCTCCTTGA TGAAGCCACC GCCTCTATGG ACTCCAAGAC TGACACCCTG 3780  
 20 GTTCAGAACCA CCATCAAAGA TGCCTTCAAG GGCTGCACAT TGCTGACCAT CGCCCACCGC 3840  
 CTCAACACAG TTCTCAACTG CGATCAGCTC CTGGTTATGG AAAATGGGAA GGTGATTGAG 3900  
 TTTGACAAGC CTGAAGTCCT TGCAGAGAAG CCAGATTCTG CATTTGCGAT GTTACTAGCA 3960  
 GCAGAAGTCA GATTGTAG

Seq ID NO: 513 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | |  
 30 MVGEGPYLIS DLDQRRRRS FAERYDPSLK TMIPVRPCAR LAPNPVDDAG LLSFATFSWL 60  
 TPVMVKGYRQ RLTVDTLPL STYDSSDTNA KRFRVLWDEE VARVGPEKAS LSHVVWKFQR 120  
 TRVLMDIVAN ILCIIMAAIG PTVLIHQILQ QTERTSGKVW VGIGLCIALF ATEFTKVFFW 180  
 ALAWAINYRT AIRLKVALLST LVFENLVSEK TLTHISVGEV LNILSSDSYS LFEAALFCPL 240  
 35 PATIPILMVF CAAYAFFILG PTALIGISVY VIFIPVQMF M AKLNSAFRRS AILVTDKRVQ 300  
 TMNEFLTCIR LIKMYAWEKS FTNTIQDIRR RERKLLKAG FVQSGNSALA PIVSTIAIVL 360  
 TLSCHILLRR KLTAPVAFSV IAMFNVMKFS IAILPFSIKA MAEANVSLRR MKKILIDKSP 420  
 PSYITQPEDP DTVLLLANAT LTWEHEASRK STPKKLQNK RHLCKKQSR AYSESPPAK 480  
 GATGPEEQSD SLKSVLHSIS FVVRKLCRYP EAQLLAWRWP AVFVGRIIRG YRPHGFSAD 540  
 40 KDESRRLLTW PQEVDRTQRA AKYLKILGI CGNVGSGKSS LLAALLGQM LQKGVVAVNG 600  
 TLAYVSQAW IFHGNVRENI LFGEKYDHQR YQHTVRVCG L QKDLNLPYG DLTEIGERGL 660  
 NLSGGQRQRI SLARAVYSR QLYLLDDPLS AVDAHVGKHV FECCIKKTLR GKTVVVLVTHQ 720  
 LQFLESCDEV ILLEDGEICE KGTHKELMEE RGRYAKLIHN LRGLQFKDPE HLYNAMVEA 780  
 FKESPAEREE DAGIIGYLLS LFTVFLFLM IGSAAFSNWW LGLWLDKGSR MTCGPQGNRT 840  
 MCEVGAVLAD IGQHVYQWVY TASMVFMVLF GVTKGFTVTK TTLMASSSLH DTVFDKILKS 900  
 45 PMSFFDTTPT GRMLNRFSD MDEL DVLRPS HAENFLQFF MVVFILVILA AVFPAVLLV 960  
 ASLAVGFIL LRIFHRGVQE LKKVENVRSS PWFTHITSSM QGLGIHAYG KKESCITYTS 1020  
 SKGLSLSYII QLSGLLQVCV RTGTETQAKF TSVELLREYI STCVPECTHP LKVGTCPKDW 1080  
 PSCGEITFRD YQMYRDNTF LVLDSLNLNI QSGQTVGIVG RTGSGKSSLG MALFRLVEPA 1140  
 50 SGTIFIDEVD ICILSLEDLR TKLTVIPQDP VLFVGTVRYN LDPFESHTDE MLWQVLERTF 1200  
 MRDTIMKLPE KLQAEVTENG ENFSVGERQL LCVARALLRN SKIILLDEAT ASMDSKTDTL 1260  
 VQNTIKDAFK GCTVLTIAHR LNTVLNCDHV LVMENGKVIE FDKPEVLAEK PDSAFAMLLA 1320  
 AEVRL

Seq ID NO: 514 DNA sequence  
 Nucleic Acid Accession #: Z31560  
 Coding sequence: 1-966

1 11 21 31 41 51  
 | | | | |  
 60 CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCCCGG CCCGAGCAA 60  
 ACTTCGGGGG GCGGCGGCGG CAACTCCACC GCGGCGGCGG CCGGCGGCAA CCAGAAAAC 120  
 AGCCCGGACC GCGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCCG CGGGCAGCGG 180  
 CGCAAGATGG CCCAGGAGAA CCCCAAGATG CACAACCTCG AGATCAGCAA GCGCCTGGGC 240  
 GCGGAGTGG AACTTTTGTG GGAGACGGAG AAGCGGCGGT TCATCGACGA GGCTAAGCGG 300  
 65 CTGCGAGCGC TGCACATGAA GGAGACCCCG GATTATAAAT ACCGCGCCCG GCGGAAAACC 360  
 AAGACGCTCA TGAGAAGGA TAAGTACACG CTGCCCCGGC GGCTGCTGGC CCGGCGGCGG 420  
 AATAGCATGG CGAGCGGGGT CGGGGTGGGC GCGGCGCTGG GCGGCGGCGT GAACAGCGC 480  
 ATGGACAGTT ACGCGCACAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540  
 CAGCTGGGCT ACCCGCAGCA CCCGGGCTC AATGCGCACG GCGCAGCGCA GATGCAGCCC 600  
 70 ATGCACCGCT ACGACGTGAG CGCCCTGCAC TACAACCTCA TGACCAGCTC GCAGACCTAC 660  
 ATGAACGGCT CGCCACCTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT 720  
 CTTGGCTCCA TGGGTTCTGGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCC TGTGGTTACC 780  
 TCTTCTCTCC ACTCCAGGGC GCCCTGCCAG GCGGGGACC TCCGGGACAT GATCAGCATG 840  
 75 TATCTCCCCG GCGCCGAGGT GCCGGAACCC GCGGCCCCCA GCAGACTTCA CATGTCCCAG 900  
 CACTACCAGA GCGGCCCCGT GCCCGGCACG GCCATTAAAC GCACACTGCC CCTCTCACAC 960  
 ATGTGAGGGC CGGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAAGAAAA ACGAGGGAAA 1020  
 TGGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCCG TACGCTCAAA 1080  
 AAAAA

Seq ID NO: 515 Protein sequence  
 Protein Accession #: CAA83435

1 11 21 31 41 51  
 | | | | |  
 85 HSARMYNMME TELKPPGPQQ TSGGGGNGST AAAAGGNQKN SPDRVKRPMN AFMVWSRGQR 60  
 RKMAQENPKM HNSEISKRLG AEWKLLSETE KRPFIDEAKR LRALHMKEHP DYKYRPRRKT 120  
 KTLMKDKYT LPGGLLAPGG NSMASGVGVG AGLGAGVNQR MDSYAHMNGW SNGSYSMMQD 180



QLGYPPQHPGL NAHGAAQMQP MHRVDVSALQ YNSMTSSQTY MNGSPTYSMS YSQQGTGPGMA 240  
 LGSMGSSVVK SASSPPVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPVPEP AAPSR LHMSQ 300  
 HYQSGPVPGT AINGTLPLSH M

5 Seq ID NO: 516 DNA sequence  
 Nucleic Acid Accession #: U91618  
 Coding sequence: 29..541

10 1 11 21 31 41 51  
 | | | | | |  
 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60  
 CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAAATGAA 120  
 AGCATTAGAA GCAGATTTCT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180  
 TCCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC 240  
 15 AGCTGAGGAA ACAGGAGAAG TTCATGAAGA GGAGCTTGTT GCAAGAAGGA AACTTCCTAC 300  
 TGCTTTAGAT GGCTTTAGCT TGGAAAGCAAT GTTGACAATA TACCAGCTCC ACAAATCTG 360  
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA 420  
 TGACAAAAT GGAAAGGAAG AAGTCATAAA GAGAAAATTT CTTATATTC TGAAACGGCA 480  
 GCTGTATGAG AATAACCCA GAAGACCTTA CATACTCAA AGAGATTCTT ACTATTACTG 540  
 20 AGAGAATAAA TCATTATTT ACATGTGATT GTGATTATC ATCCCTAAT TAAATATCAA 600  
 ATTATATTTG TGTGAAAATG TGACAAACAT ACTTATCTGT CTCTCTACA ATTGTGGTTT 660  
 ATTTGAATGT TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAA TAAATCTAAA 720  
 TCTTCAAAA AAAAAAAAAA AAATGGGGCC GCAATT

25 Seq ID NO: 517 Protein sequence  
 Protein Accession #: AAB50564

30 1 11 21 31 41 51  
 | | | | | |  
 MMAGMKIQLV CMLLLAFSSW SLCS DSEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60  
 VCSLVNNLNS PAETGEVHE EELVARRKLP TALDGFSL EAMLTIIQLHKI CHSRAFQHWE 120  
 LIQEDILDTG NDKNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSYYY

35 Seq ID NO: 518 DNA sequence  
 Nucleic Acid Accession #: NM\_006536.2  
 Coding sequence: 109..2940

40 1 11 21 31 41 51  
 | | | | | |  
 ACCTAAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60  
 ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120  
 AGCATTGCAG GTCCTATTG CAACCTGAAG TTTGTGACTC TCCTGGTTGC CTTAAGTTCA 180  
 GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC 240  
 45 ATTGCAATTA ATCCTCAGGT ACCTGAGAAT CAGAACCTCA TCTCAAACAT TAAGGAAATG 300  
 ATAAGTGAAG CTTCAATTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTTCAAGAA 360  
 ATAAAGATTT TAATACCTGC CACATGGAAT GCTAATAATA ACAGCAAAAT AAAACAAGAA 420  
 TCAATGAAA AGGCAATGT CATAGTGACT CACTGGTATG GGGCACATGG AGATGATCCA 480  
 TACACCCTAC AATACAGAGG GTGTGGAAAA GAGGGAAAT ACATTCATTT CACACCTAAT 540  
 50 TTCTACTGA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCCGAGT GTTTGTCCAT 600  
 GAATGGGCCC ACCTCCGTTG GGGTGTGTTT GATGAGTATA ACAATGACAA ACCTTTCTAC 660  
 ATAAATGGGC AAAATCAAAAT TAAAGTGACA AGGTGTTTAT CTGACATCAC AGGCATTTTT 720  
 GTGTGTGAAA AAGGTCCCTG CCCCCAAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780  
 GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840  
 55 AGTTTATCTT CTGTGGTTGA ATTTTGTAAAT GCAAGTACCC ACAACCAAGA AGCACCAAAC 900  
 CTACAGAACC AGATGTGCAG CCTCAGAAGT GCATGGGATG TAATCACAGA CTCTGCTGAC 960  
 TTTTACCACA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCTCCACAT ATTCTCGCTT 1020  
 GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTGCTGGATG TGTCCAGCAA GATGGCAGAG 1080  
 GCTGACAGAC TCCTTCAACT ACAACAAGCC GCAGAATTTT ATTTGATGCA GATTGTTGAA 1140  
 60 ATTCATACCT TCGTGGGCAT TGCCAGTTTC GACAGCAAAG GAGAGATCAG AGCCAGCTA 1200  
 CACCAATTA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260  
 TCAGCTAAAA CAGACATCAG CATTTGTTCA GGGCTTAAGA AAGGATTGTA GGTGGTTGAA 1320  
 AAAGTGAATG GAAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG 1380  
 CTTCTTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATFCA CTCCATTGCC 1440  
 CTGGGTTTAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500  
 65 TTCTTTGTTT CAGATATATC AAATCCAATC AGACTGATTG ATGCTTTCAG TAGAATTTCC 1560  
 TCTGGAACTG GAGACATTTT CCAGCAACAT ATTCAGCTTG AAAGTACAGG TGAAAATGTC 1620  
 AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680  
 ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCCTGAGA TTATATATAT TGATCCTGAT 1740  
 70 GGACGAAAAT ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800  
 TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCCTGAACAA TACCCATCAT 1860  
 TCTCTGCAAG CCTGAAAAT GACAGTGACC TCTCGCGCCT CCAACTCAGC TGTGCCCCCA 1920  
 GCCACTGTGG AAGCCTTTGT GGAAAGAGAC AGCCTCCATT TTCTCATCC TGTGATGATT 1980  
 TATGCCAATG TGAACAGGG ATTTTATCCC ATTCTTAATG CCACTGTCAC TGCCACAGTT 2040  
 75 GAGCCAGAGA CTGGAGATCC TGTTACGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100  
 GTTATAAAAA ATGATGGAAT TTAATCGAGG TATTTTTTCT CTTTGTCTGC AAATGGTAGA 2160  
 TATAGCTTGA AAGTGCATGT CAATCACTCT CCCAGCATAA GCACCCAGC CCACTCTATT 2220  
 CCAGGGAGTC ATGCTATGTA TGTACCAGGT TACACAGCAA ACGGTAATAT TCAGATGAAT 2280  
 GCTCCAAGGA AATCAGTAG CAGAAATGAG GAGAGCGAA AGTGGGCTT TAGCCGAGTC 2340  
 AGCTCAGGAG GCTCCTTTTC AGTGCTGGGA GTTCCAGCTG GCCCCACCC TGATGTGTTT 2400  
 80 CCACCATGCA AATTATTGA CCTGGAAGCT GTAAAAGTAG AAGAGGAATT GACCCTATCT 2460  
 TGGACAGCAC CTGGAGAAGA CTTTGATCAG GGCCAGGCTA CAAGCTATGA AATAAGAATG 2520  
 AGTAAAAGTC TACAGAAAT CCAAGATGAC TTAAACAATG CTATTTTAGT AAATACATCA 2580  
 AAGCGAAATC CTCAGCAAGC TGGCATCAGG GAGATATTTA CGTTCTCACC CCAGATTTC 2640  
 85 ACGAATGGAC CTGAACATCA GCCAAATGGA GAAACACATG AAAGCCACAG AATTTATGTT 2700  
 GCAATACGAG CAATGGATAG GAATCCTTA CAGTCTGCTG TATCTAACAT TGCCAGGCG 2760  
 CCTCTGTTTA TTCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820  
 GGAGTTTTAA CAGCAATGGG TTTGATAGGA ATCATTTGCC TTATTATAGT TGTGACACAT 2880



CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATAA 2940  
 ATAAATATCC AAAGTGCTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAAA 3000  
 CATACTAACA AAGTCAAAAT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060  
 ATACAGATAA GATTTTACACA TGGTAGATCA ACAATTCTTT TTGGGGGTAG ATTAGAAAAAC 3120  
 CCTTACACTT TGGCTATGAA CAAATAATAA AAATTATTCT TTAAAGTAAT GTCTTTAAAG 3180  
 GCAAAGGGAA GGGTAAAGTC GGACCAGTGT CAAGGAAAGT TTGTTTTATT GAGGTGGAAA 3240  
 AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA 3300  
 TCATTTAGTT ACTTTGATTA ATTTTCTTTT TCTCCTTATC TGTGCAGTAC AGGTTGCTTG 3360  
 TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAAG CTCTTTACCT 3420  
 CTTGCTATTT TGTATATAT ATTTTCTAGT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480  
 TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540  
 TTTATGACAA AGGTCTATTG AATTTATTTG TGTGTAAGTT TCTACTCCCA TCAAAGCAGC 3600  
 TTTCTAAGTT TATTGCCTTG GGTATTATG GAATGATAGT TATAGCCCN TATAATGCCT 3660  
 TACCTAGGAA A

Seq ID NO: 519 Protein sequence  
 Protein Accession #: NP\_006527.1

1 11 21 31 41 51  
 | | | | |  
 MTQRSIAGPI CNLKFTVLLV ALSSELPLFLG AGVQLQDNGY NGLLIAINPQ VPENQNLISN 60  
 IKEMITEASF YLEFNATKRRV FFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAAH 120  
 GDDPYTLQYR GCGKEGKYIH FTPNFLLNDN LTAGYGSRRG VFWHEWAHLR WGVFDEYNND 180  
 KPFYINGQNR IKVTRCSSDI TGIFVCEKGP CPQENCIISK LFKEGCTFIY NSTQNATASI 240  
 MFMQSLSSVV EFCNASTHNQ EAPNLQNMCM SLRSAWDVIT DSADFHHSFP MNGTELPPPP 300  
 TFSLVQAGDK VVCLVLDVSS KMAEADRLLO LQQAEEFYLM QIVEIHTFVG IASFDSKGEI 360  
 RAQLHQINSN DDRKLLVSYL PTTVSAKTDI SICSGLKKGF EVVEKLNGKA YGSVMILVTS 420  
 GDDKLLGNCL PTVLSSGSTI HSIALGSSAA PNLEELSRLT GGLKFFVPDI SNSNSMIDAF 480  
 SRISSTGTDI FQHQIQLST GENVKPHHQL KNTVTVDNTV GNDTMFLVTW QASGPPEIIL 540  
 FDPDGRKYIT NNFITNLTFR TASLWIPGTA KPGHWYTLN NTHHSLQALK VTVTSRASNS 600  
 AVPPATVEAF VERDSLHFPV PVMIIYANVKQ GFYPILNATV TATVEPETGD PVTLRLLDDG 660  
 AGADVINKNDG IYSRYFFSFA ANGRYSLKVH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720  
 IQMNAPRKSV GRNEEERKVG FSRVSSGGSF SVLGVFAGPH PDVFPCKII DLEAVKVEEE 780  
 LTLSTWAPGE DFDQQAQTSY EIRMSKSLQN IQDDFNAIL VNTSKRNPQQ AGIREIFTFS 840  
 PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900  
 LILKGVLTM GLIGIICLII VVTHHTLSRK KRADKKENG T KLL

Seq ID NO: 520 DNA sequence  
 Nucleic Acid Accession #: NM\_000228.1  
 Coding sequence: 82..3600

1 11 21 31 41 51  
 | | | | |  
 GCTTTCAGGC GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAAGG TCCTTCTCTGG 60  
 GGATCACCCC ATTGGCTGAA GATGAGACCA TTCTTCTCTT TGTGTTTTCG CCTGCCTGGC 120  
 CTCCTGCATG CCCAACAGC CTGCTCCCGT GGGGCCTGCT ATCCACCTGT TGGGGACCTG 180  
 CTTGTTGGGA GGACCCGTT TCTCCGACT TCATCTACCT GTGGACTGAC CAAGCCTGAG 240  
 ACCTACTGCA CCCAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCCAGGCAG 300  
 CCTCACAACT ACTACAGTCA CCGAGTAGAG AATGTGGCTT CATCCTCCGG CCCCATGCGC 360  
 TGGTGGCAGT CCCAGAAATGA TGTGAACCTT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420  
 TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGCCCGCCGG CATGCTGATT 480  
 GAGCGCTCCT CAGACTTCGG TAAGACCTGT CGAGTGTACC AGTACCTGGC TGCCGACTGC 540  
 ACCTCCACCT TCCCTCGGGT CCGCCAGGT GCGCCTCAGA GCTGGCAGGA TGTTCGGTGC 600  
 CAGTCCCTGC CTCAGAGGCC TAATGCACGC CTAAATGGGG GGAAGGTCCA ACTTAACCTT 660  
 ATGGATTTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAA AAATCAAGA GGTGGGGGAG 720  
 ATCACAACT TGAGAGTCAA TTTCACCAGG CTGGCCCCCTG TGCCCCAAAG GGGCTACCAC 780  
 CCTCCCAGCG CTTACTATGC TGTGTCCCAG CTCCGTCTGC AGGGGAGCTG CTTCTGTCTC 840  
 GGCCATGCTG ATCGCTGCGC ACCCAAGCCT GGGGCTCTG CAGGCCCCCTC CACCGCTGTG 900  
 CAGGTCCACG ATGTCTGTGT CTGCTGACAC AACACTGCCG GCCCAAATTG TGAGCGCTGT 960  
 GCACCCCTCT ACACAAACCG GCCCTGGAGA CCGCGGAGG GCCAGGACGC CCATGAATGC 1020  
 CAAAGGTGCG ACTGCAATGG GCACTCAGAG ACATGTCACT TTGACCCCGC TGTGTTTGCC 1080  
 GCCAGCCAGG GGGCATATGG AGGTGTGTGT GACAATTGCC GGGACCACAC CGAAGGCAAG 1140  
 AACTGTGAGC GGTGTGAGCT GCACTATTTT CGGAACCGGC GCCCGGGAGC TTCCATTGAG 1200  
 GAGACCTGCA TCTCTGCGA GTGTGATCCG GATGGGGCAG TGCCAGGGGC TCCCTGTGAC 1260  
 CCAGTGACCG GGCAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320  
 AAGCCGGGCT TCACTGGACT CACCTACGCC AACCCGCAGG GCTGCCACCG CTGTGACTGC 1380  
 AACATCTGG GGTCCCGGAG GGACATGCCG TGTGACGAGG AGAGTGGGCG CTGCCTTTGT 1440  
 CTGCCCCAAG TGGTGGGTCC CAAATGTGAC CAGTGTGCTC CCTACCACTG GAAGCTGGCC 1500  
 AGTGGCCAGG GCTGTGAACC GTGTGCCTGC GACCCGCACA ACTCCCTCA GCCCACAGTG 1560  
 CAACAGTTT ACAGGGCAGT GCCCTGTGCG GAAGGCTTTG GTGGCCTGAT GTGCAGCGCT 1620  
 GCAGCCATCC GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAGG ATGCCGAGCC 1680  
 TGTGACTGTG ATTTCCGGGG AACAGAGGCG CCGGCTGCG ACAAGGCATC AGGCCGCTGC 1740  
 CTCTGCCGCC CTGGCTTGAC CGGGCCCCGC TGTGACAGT GCCAGCGAGG CTTACTGCAAT 1800  
 CGCTACCCGG TGTGCGTGGC CTGCCACCCT TGCTTCCAGA CCTATGATGC GGACCTCCGG 1860  
 GAGCAGGCCC TGCGCTTTGG TAGACTCCGC AATGCCACCG CCAGCCTGTG GTCAGGGCCT 1920  
 GGGCTGGAGG ACCGTGGCCT GGCCTCCCGG ATCCTAGATG CAAAGAGTAA GATTGAGCAG 1980  
 ATCCGAGCAG TTCTCAGCAG CCCCAGCATC ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT 2040  
 GCCATCCTCT CCTCAGGCG AACTCTCCAG GGCCTGCAGC TGGATCTGCC CCTGGAGGAG 2100  
 GAGACGTTGT CCCTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCCTT 2160  
 ACTATGTATC AGAGGAAGAG GGAGCAGTTT GAAAAATAA GCAGTGCTGA TCCTTCAGGA 2220  
 GCCTTCCGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGCTCA GCAGGTCTCC 2280  
 GACAGCTCGC GCCTTTTGGA CCAGCTCAGG GACAGCCGGA GAGAGGCAGA GAGGCTGGTG 2340  
 CGGCAGGCGG GAGGAGGAGG AGGCACCGG AGCCCCAAGC TTGTGGCCCT GAGGCTGGAG 2400  
 ATGTCTTCGT TGCTTGACCT GACACCCACC TTCAACAAGC TCTGTGGCAA CTCCAGGCAG 2460  
 ATGGCTTGCA CCCCATATC ATGCCCTGGT GAGCTATGTC CCCAAGACAA TGGCACAGCC 2520  
 TGTGGCTCCC GCTGCAGGGG TGTCTTCCC AGGGCCGGTG GGGCCTTCTT GATGGCGGGG 2580  
 CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCAGCTCC AGCGGACCAG GCAGATGATT 2640



	AGGGCAGCCG	AGGAATCTGC	CTCACAGATT	CAATCCAGTG	CCCAGCGCTT	GGAGACCCAG	2700
	GTGAGCGCCA	GCCGCTCCCA	GATGGAGGAA	GATGTCAGAC	GCACACGGCT	CCTAATCCAG	2760
	CAGGTCCGGG	ACTTCCTAAC	AGACCCCGAC	ACTGATGCAG	CCACTATCCA	GGAGGTCAGC	2820
5	GAGGCCGTGC	TGGCCCTGTG	GCTGCCACAC	GACTCAGCTA	CTGTTCTGCA	GAAGATGAAT	2880
	GAGATCCAGG	CCATTGCAGC	CAGGCTCCCC	AACGTGGACT	TGGTGCTGTC	CCAGACCAAG	2940
	CAGGACATTG	CGCGTGCCCG	CCGTTGTCAG	GCTGAGGCTG	AGGAAGCCAG	GAGCCGAGCC	3000
	CATGCAGTGG	AGGGCCAGGT	GGAAGATGTG	GTTGGGAACC	TGCGGCAGGG	GACAGTGGCA	3060
	CTGCAGGAAG	CTCAGGACAC	CATGCAAGGC	ACCAGCCGCT	CCCTTCGGCT	TATCCAGGAC	3120
10	AGGGTTGCTG	AGGTTTCAGCA	GGTACTGCGG	CCAGCAGAAA	AGCTGGTGAC	AAGCATGACC	3180
	AAGCAGCTGG	GTGACTTCTG	GACACGGATG	GAGGAGCTCC	GCCACCAAGC	CCGGCAGCAG	3240
	GGGGCAGAGG	CAGTCCAGGC	CCAGCAGCTT	GCGGAAGGTG	CCAGCGAGCA	GGCATTGAGT	3300
	GCCCAAGAGG	GATTTGAGAG	AATAAAACAA	AAGTATGCTG	AGTTGAAGGA	CCGGTTGGGT	3360
	CAGAGTTCCA	TGCTGGGTGA	GCAGGGTGCC	CGGATCCAGA	GTGTGAAGAC	AGAGGCAGAG	3420
15	GAGCTGTTTG	GGGAGACCAT	GGAGATGATG	GACAGGATGA	AAGACATGGA	GTTGGAGCTG	3480
	CTGCGGGGCA	GCCAGGCCAT	CATGCTGCGC	TGCGCGGACC	TGACAGGACT	GGAGAAGCGT	3540
	GTGGAGCAGA	TCCGTGACCA	CATCAATGGG	CGCGTGCTCT	ACTATGCCAC	CTGCAAGTGA	3600
	TGCTACAGCT	TCCAGCCCGT	TGCCCCACTC	ATCTGCCGCC	TTTGCTTTTG	GTTGGGGGCA	3660
	GATTGGGTTG	GAATGCTTTC	CATCTCCAGG	AGACTTTCAT	GCAGCCTAAA	GTACAGCCTG	3720
20	GACCACCCCT	GGTGTGTAGC	TAGTAAGATT	ACCCTGAGCT	GCAGCTGAGC	CTGAGCCAAT	3780
	GGGACAGTTA	CACCTGACAG	ACAAAGATGG	TGGAGATTGG	CATGCCATTG	AAACTAAGAG	3840
	CTCTCAAGTC	AAGGAAGCTG	GGCTGGGATG	TATCCCCCGC	CTTTAGTTCT	CCACTGGGGA	3900
	GGAATCCTGG	ACCAAGCACA	AAAACCTAAC	AAAAGTGATG	TAAAAATGAA	AAGCCAAATA	3960
	AAAATCTTTG	G					

25 Seq ID NO: 521 Protein sequence  
Protein Accession #: NP\_000219.1

	1	11	21	31	41	51	
30	MRPFFLLCFA	LPGLLHAQQA	CSRGACYPPV	GDLLVGRTRF	LRASSTCGLT	KPETYCTQYG	60
	EWQMKCKCKD	SRQPHNYSH	RVENVASSSG	PMRWQSQND	VNPVSLQLDL	DRRFQLQEVN	120
	MEFQGPMPAG	MLIERSDFG	KTWRVYQYLA	ADCTSTFPRV	RQGRPQSWQD	VRCQSLPQRP	180
	NARLNGGKVQ	LNLMDLVSGI	PATQSQKIQE	VGEITNLRVN	FTRLAPVPQR	GYHPPSAYYA	240
35	VSQRLRLQGSC	FCHGHADRC	PKPGASAGPS	TAVGVHDVCV	CQHNTAGPNC	ERCAPFYNNR	300
	PWRPAEGQDA	HECQRCDNCG	HSETCHFDPA	VFAASQGAYG	GVCNDCRDHT	EGKNCERCQL	360
	HYFRNRRPGA	SIQETCISCE	CDPDGAVPGA	PCDPVTGQCV	CKEHVQGERC	DLCKPGFTGL	420
	TYANPQGCHR	CDCNILGSRR	DMPDDEESGR	CLCLPNVVG	KCDQCAPYHW	KLASGQGCEP	480
	CACDPHNSPQ	PTVQPVHRAV	PCREGFGGLM	CSAAAIRQCP	DRYGDVATG	CRACDCDFRG	540
40	TEGPGCDKAS	GRCLCRPGLT	GPRCDQCQRG	YCNRYPVCSA	CHPCFQTYDA	DLREQALRFG	600
	RLRNATASLW	SGPGLDRGL	ASRILDAKSK	IEQIRAVLSS	PAVTEQEVQA	VASAILSLRR	660
	TLQGLQLDLP	LEEETLSLPR	DLESIDRSFN	GLLTMYQRKR	EQFEKISSAD	PSGAFRMLST	720
	AYEQSAQAAQ	QVSDSSRLLD	QLRDSRREAE	RLVRQAGGGG	GTGSPKLVAL	RLEMSSLPDL	780
	TPTFNKLCGN	SRQMACTPIS	CPGELCPQDN	GTACGSRCRG	VLPRAGGAFI	MAGQVAEQLR	840
45	GFNAQLQRTR	QMIRAAEESA	SQIQSSAQRL	ETQVSASRSQ	MEEDVRRTRL	LIQQVRDFLT	900
	DPDIDAATIQ	EVSEAVLALW	LPTDSATVLR	KMNEIQAIQA	RLPNVDLVLS	QTKQDIARAR	960
	RLQAEAEEAR	SRAHAVEGQV	EDVVSNTLRQ	TVALEAQTDT	MQGTSRSLRL	IQDRVAEVQQ	1020
	VLRPAEKLVT	SMTKQLGDFW	TRMEELRHQA	RQQAEEAVQA	QQLAEGASEQ	ALSAQEGFER	1080
50	IKQKYAELKD	RLGQSSMLGE	QGARIQSVKT	EAEELFGETM	EMMDRMKDME	LELLRGSQAI	1140
	MLRSADLTGL	EKRVEQIRDH	INGRVLYYAT	CK			

Seq ID NO: 522 DNA sequence  
Nucleic Acid Accession #: NM\_001944.1  
Coding sequence: 84..3083

55	1	11	21	31	41	51	
	TTTTCTTAGA	CATTAAGTGC	AGACGGCTGG	CAGGATAGAA	GCAGCGGCTC	ACTTGGAATT	60
	TTTCACCAGG	GAAATCAGAG	ACAATGATAG	GGCTCTTCCC	CAGAACTACA	GGGGCTCTGG	120
60	CCATCTTCGT	GGTGGTCATA	TGTTGTCATG	GAGAATTGCG	AATAGAGACT	AAAGGTCAAT	180
	ATGATGAAGA	AGAGATGACT	ATGCAACAAG	CTAAAAGAAG	GCAAAAACGT	GAATGGGTGA	240
	AATTTGCCAA	ACCCTGCAGA	GAAGGAGAAG	ATAACTCAAA	AAGAAACCCA	ATTGCCAAGA	300
	TTACTTCAGA	TTACCAAGCA	ACCCAGAAAA	TCACCTACCG	AATCTCTGGA	GTGGGAATCG	360
	ATCAGCCGCC	TTTTGGAATC	TTTGTGTGTT	ACAAAAACAC	TGGAGATATT	AACATAACAG	420
65	CTATAGTCGA	CCGGGAGGAA	ACTCCAAGCT	TCCTGATCAC	ATGTCGGGCT	CTAAATGCCC	480
	AAGGACTAGA	TGTAGAGAAA	CCACTTATAC	TAACGGTTAA	AATTTTGGAT	ATTAATGATA	540
	ATCCTCCAGT	ATTTTTCACAA	CAAATTTTCA	TGGGTGAAAT	TGAAGAAAAT	AGTGCCTCAA	600
	ACTCACTGGT	GATGATACTA	AATGCCACAG	ATGCAGATGA	ACCAAACAC	TTGAATTCTA	660
	AAATTGCCTT	CAAAATGTCT	TCTCAGGAAC	CAGCAGGCAC	ACCCATGTTC	CTCCTAAGCA	720
70	GAAACACTGG	GGAAGTCCGT	ACTTTGACCA	ATTCCTTTGA	CCGAGAGCAA	GCTAGCAGCT	780
	ATCGTCTGGT	TGTGAGTGGT	GCAGACAAAG	ATGGAGAAGG	ACTATCAACT	CAATGTGAAT	840
	GTAATATTAA	AGTGAAAGAT	GTCAACAGATA	ACTTCCCAAT	GTTTAGAGAC	TCTCAGTATT	900
	CAGCACGTAT	TGAAGAAAAAT	ATTTTAAGTT	CTGAATTACT	TCGATTTCAA	GTAACAGATT	960
	TGGATGAAGA	GTACACAGAT	AATTGGCTTG	CAGTATATTT	CTTTACCTCT	GGGAATGAAG	1020
75	GAAATTGGTT	TGAAATACAA	ACTGATCCTA	GAACATAATGA	AGGCATCCTG	AAAGTGGTGA	1080
	AGGCTCTAGA	TTATGAACAA	CTACAAAGCG	TGAAACTTAG	TATTGCTGTC	AAAAACAAAG	1140
	CTGAATTTCA	CCAATCAGTT	ATCTCTCGAT	ACCGAGTTCA	GTCAACCCCA	GTCACAATTC	1200
	AGGTAATAAA	TGTAAGAGAA	GGAATTGCAT	TCCGTCCTGC	TTCCAAGACA	TTTACTGTGC	1260
80	AAAAAGGCAT	AAGTAGCAAA	AAATTGGTGG	ATTATATCCT	GGGAACATAT	CAAGCCATCG	1320
	ATGAGGACAC	TAACAAAGCT	GCCTCAAATG	TCAAATATGT	CATGGGACGT	AACGATGGTG	1380
	GATACCTAAT	GATTGATTCA	AAAACCTGCT	AAATCAAATT	TGTCAAAAAT	ATGAACCGAG	1440
	ATTCTACTTT	CATAGTTAAC	AAAACAATCA	CAGCTGAGGT	TCTGGCCATA	GATGAATACA	1500
	CGGGTAAAC	TTCTACAGGC	ACGGTATATG	TTAGAGTACC	CGATTTCAAT	GACAATTGTC	1560
	CAACAGCTGT	CCTCGAAAAA	GATGCAGTTT	GCAGTTCTTC	ACCTTCCGTG	GTTGTCTCCG	1620
85	CTAGAACACT	GAATAATAGA	TACACTGGCC	CAGTATACATT	TGCACTGGAA	GATCAACCTG	1680
	TAAAGTTGCC	TGCCGTATGG	AGTATCACAA	CCCTCAATGC	TACCTCGGCC	CTCCTCAGAG	1740
	CCCAGGAACA	GATACCTCCT	GGAGTATACC	ACATCTCCCT	GGTACTTACA	GACAGTCAGA	1800
	ACAATCGGTG	TGAGATGCCA	CGCAGCTTGA	CACTGGAAGT	CTGTCAGTGT	GACAACAGGG	1860



	GCATCTGTGG	AACTTCTTAC	CCAACCACAA	GCCCTGGGAC	CAGGTATGGC	AGGCCGCACT	1920
	CAGGGAGGCT	GGGGCCTGCC	GCCATCGGCC	TGCTGCTCCT	TGGTCTCCTG	CTGCTGCTGT	1980
	TGGCCCCCCT	TCTGCTGTTG	ACCTGTGACT	GTGGGGCAGG	TTCTACTGGG	GGAGTGACAG	2040
5	GTGGTTTTAT	CCCAGTTCCT	GATGGCTCAG	AAGGAACAAT	TCATCAGTGG	GGAATTGAAG	2100
	GAGCCCATCC	TGAAGACAAG	GAAATCACAA	ATATTTGTGT	GCCTCCTGTA	ACAGCCAATG	2160
	GAGCCGATTT	CATGGAAAGT	TCTGAAGTTT	GTACAAATAC	GTATGCCAGA	GGCACAGCGG	2220
	TGGAAGGCAC	TTCAGGAATG	GAAATGACCA	CTAAGCTTGG	AGCAGCCACT	GAATCTGGAG	2280
	GTGCTGCAGG	CTTTGCAACA	GGGACAGTGT	CAGGAGCTGC	TTCAGGATTC	GGAGCAGCCA	2340
10	CTGGAGTTGG	CATCTGTTCC	TCAGGGCAGT	CTGGAACCAT	GAGAACAAGG	CATTCCACTG	2400
	GAGGAACCAA	TAAGGACTAC	GCTGATGGGG	CGATAAGCAT	GAATTTTCTG	GACTCCTACT	2460
	TTTCTCAGAA	AGCATTGACC	TGTGCGGAGG	AAGACGATGG	CCAGGAAGCA	AATGACTGCT	2520
	TGTTGATCTA	TGATAATGAA	GGCGCAGATG	CCACTGGTTC	TCCTGTGGGC	TCCGTGGGTT	2580
	GTTGCAGTTT	TATTGCTGAT	GACCTGGATG	ACAGCTTCTT	GGACTCACTT	GGACCCAAAT	2640
	TTAAAAAAT	TGCAGAGATA	AGCCTTGGTG	TTGATGGTGA	AGGCAAAGAA	GTTGAGCCAC	2700
15	CCTCTAAGA	CAGCGTTAT	GGGATTGAAT	CCTGTGGCCA	TCCCATAGAA	GTCCAGCAGA	2760
	CAGGATTTGT	TAAGTGCCAG	ACTTTGTCAG	GAAGTCAAGG	AGCTTCTGCT	TTGTCCGCCT	2820
	CTGGGTCTGT	CCAGCCAGCT	GTTTCCATCC	CTGACCCTCT	GCAGCATGGT	AACTATTTAG	2880
	TAACGGAGAC	TTACTCGGCT	TCTGGTTCCC	TCGTGCAACC	TTCCACTGCA	GGCTTTGATC	2940
20	CACTTCTCAC	ACAAAATGTG	ATAGTGACAG	AAAGGGTGAT	CTGTCCCATT	TCCAGTGTTC	3000
	CTGGCAACCT	AGCTGGCCCA	ACGCAGCTAC	GAGGGTCACA	TACTATGCTC	TGTACAGAGG	3060
	ATCCTTGCTC	CCGTCTAATA	TGACCAGAAT	GAGCTGGAAT	ACCACACTGA	CCAAATCTGG	3120
	ATCTTTGGAC	TAAAGTATTC	AAAATAGCAT	AGCAAAGCTC	ACTGTATTGG	GCTAATAATT	3180
	TGGCACTTAT	TAGCTTCTCT	CATAAACTGA	TCACGATTAT	AAATTAAATG	TTTGGGTTCA	3240
25	TACCCCAAAA	GCAATATGTT	GTCACTCCTA	ATTCTCAAGT	ACTATTCAAA	TTGTAGTAAA	3300
	TCTTAAAGTT	TTTCAAAACC	CTAAAAATCAT	ATTCGC			

Seq ID NO: 523 Protein sequence  
Protein Accession #: NP\_001935.1

30	1	11	21	31	41	51	
	MMGLFPRTTG	ALAIFVVVIL	VHGELEIETK	GQYDEEEMTM	QQAARRQKRE	WVKFAKPCRE	60
	GEDNSKRNP	AKITSDYQAT	QKITYRISGV	GIDQPPFGIF	VVDKNTGDIN	ITAIVDREET	120
35	PSFLITCRAL	NAQGLDVEKP	LILTVKILDI	NDNPPVFSQQ	IFMGEIEENS	ASNSLVMILN	180
	ATDADEPNHL	NSKIAFKIVS	QEPAGTPMFL	LSRNTGEVRT	LTNSLDREQA	SSYRLVVSGA	240
	DKDGEGLSTQ	CECNIVKDV	NDNPFMERDS	QYSARIEENI	LSSELLRFQV	TDLDEEYTDN	300
	WLAVYFFTS	NEGNWFEIQT	DPRTNIGILK	VVKALDYEQL	QSVKLSIAVK	NKAEFHQSVI	360
	SRYRVQSTPV	TIQVINVREG	IAFRPASKTF	TVQKGISSKK	LVDYILGTQY	AIDEDTNKAA	420
	SNVKYVMGRN	DGGYLMIDSK	TAEIKFVKMN	NRDSTFIVNK	TITAEVLAI	EYTGKTSTGT	480
40	VYVRVPDFND	NCPTAVLEKD	AVCSSPSVSV	VSARTLNNRY	TGPYTFAL	QPVKLPAVWS	540
	ITTLNATSAL	LRAQEQIPPG	VYHISLVLT	SQNNRCEMPR	SLTLEVCQCD	NRGICGTSYP	600
	TTSPGTRYGR	PHSGRLGPAA	IGLLLLLGLL	LLAPLALLLT	CDCGAGSTGG	VTGGFIPVPD	660
	GSEGTIHQWG	IEGAHPEDKE	ITNICVPPVT	ANGADFMESS	EVCTNTYARG	TAVEGTSGME	720
45	MTTKLGAATE	SGGAAGFATG	TVSGAASGFG	AATGVGICSS	GQSGTMRTRH	STGGTNKDYA	780
	DGAISMNFLD	SYFSQKAFAC	AEEDDQGEAN	DCLLIYDNEG	ADATGSPVGS	VGCCSFIADD	840
	LDDSFLDSL	PKFKKLAEIS	LGVDGEGKEV	LPDSKDSGYG	IESCGHPIEV	QQTGFVKCQT	900
	LSGSQGSAL	SASGSVQPAV	SIPDPLQHGN	YLVETETYSAS	GSLVQPSTAG	FDPLLTQNV	960
	VTERVICPIS	SVPGNLAGPT	QLRGSHTMLC	TEDPCSRLLI			

50 Seq ID NO: 524 DNA sequence  
Nucleic Acid Accession #: XM\_058069.2  
Coding sequence: 1..1413

55	1	11	21	31	41	51	
	ATGAAGTTTC	TTCTAATACT	GCTCCTGCAG	GCCACTGCTT	CTGGAGCTCT	TCCCCTGAAC	60
	AGCTCTACAA	GCCTGGAAAA	AAATAATGTG	CTATTTGGTG	AAAGATACTT	AGAAAAATTT	120
	TATGGCCTTG	AGATAAACAA	ACTTCCAGTG	ACAAAAATGA	AATATAGTGG	AACTTAATG	180
60	AAGGAAAAAA	TCCAAGAAAT	GCAGCACTTC	TTGGGTCTGA	AAGTGACCGG	GCAACTGGAC	240
	ACATCTACCC	TGGAGATGAT	GCACGCACCT	CGATGTGGAG	TCCCCGATGT	CCATCATTTT	300
	AGGGAAATGC	CAGGGGGGCC	CGTATGGAGG	AAACATTATA	TCACCTACAG	AATCAATAAT	360
	TACACACCTG	ACATGAACCG	TGAGGATGTT	GACTACGCAA	TCCGGAAAGC	TTTCCAAGTA	420
	TGGAGTAATG	TTACCCCTTT	GAAATTCAGC	AAGATTAACA	CAGGCATGGC	TGACATTTTG	480
	GTGGTTTTTG	CCCCTGGAGC	TCATGGAGAG	TTCCATGCTT	TTGATGGCAA	AGGTGGAATC	540
65	CTAGCCCATG	CTTTTGGACC	TGGATCTGGC	ATTGGAGGGG	ATGCACATTT	CGATGAGGAC	600
	GAATTCTGGA	CTACACATTC	AGGAGGCACA	AACTTGTTCC	TCAGTCTGCT	TCACGAGATT	660
	GGCCATTCTT	TAGGTCTTGG	CCATTCTAGT	GATCCAAAGG	CCGTAATGTT	CCCCACCTAC	720
	AAATATGTTG	ACATCAACAC	ATTTTCGCCT	TCTGCTGATG	ACATACGTGG	CATTCACTCC	780
70	CTGTATGGAG	ACCCAAAAGA	GAACCAACCG	TTGCCAAATC	CTGACAAATC	AGAACCAGCT	840
	CTCTGTGACC	CCAATTTGAG	TTTTGATGCT	GTCACTACCG	TGGGAAATAA	GATCTTTTTC	900
	TTCAAAGACA	GGTCTTCTCT	GCTGAAGGTT	TCTGAGAGAC	CAAAGACCAG	TGTTAATTTA	960
	ATTTCTTCC	TATGGCCAAC	CTTGCCATCT	GCGATTGAAG	CTGCTTATGA	AATTGAAGCC	1020
	AGAAATCAAG	TTTTTCTTTT	TAAAGATGAC	AAATACTGGT	TAATTAGCAA	TTTAAGACCA	1080
75	GAGCCAAATT	ATCCCAAGAG	CATACATTCT	TTTGGTTTTT	CTAAGTTTGT	GAAAAAATTT	1140
	GATGCAGCTG	TTTTTAACCC	ACGTTTTTAT	AGGACCTACT	TCTTTGTAGA	TAACCAAGTAT	1200
	TGGAGGTATG	ATGAAAGGAG	ACAGATGATG	GACCCTGGTT	ATCCCAAAC	GATTACCAAG	1260
	AACTTCCAAG	GAATCGGGCC	TAAAATTGAT	GCAGTCTTCT	ACTCTAAAAA	CAAATACTAC	1320
	TATTTCTTCC	AAGGATCTAA	CCAATTTGAA	TATGACTTCC	TACTCCAACG	TATCACCACAA	1380
80	ACACTGAAAA	GCAATAGCTG	GTTTGGTTGT	TGA			

Seq ID NO: 525 Protein sequence  
Protein Accession #: P39900

85	1	11	21	31	41	51	
	MKFLILLLLQ	ATASGALPLN	SSTSLEKNNV	LFGERYLEKF	YGLEINKLPV	TKMKYSGNLM	60
	KEKIQEMQHF	LGLKVTGQLD	TSTLEMMHAP	RCGVDPDVHFF	REMPGGPVWR	KHYITYRINN	120



YTPDMNREDV DYAIRKAFQV WSNVTPLKFS KINTGMADIL VVFARGAHGD FHAFDGKGGI 180  
LAHAFGPGSG IGGDAHFDED EFWTTHSGGT NLFLTAVHEI GHSLGLGHSS DPKAVMFPTY 240  
KYVDINTFRL SADDIRGIQS LYGDPKENQR LPNPDNSEPA LCDPNLSFDA VTTVGNKIFF 300  
FKDRFFWLKV SERPKTSVNL ISSLWPTLPS GIEAAYEIEA RNQVFLFKDD KYWLISNLRP 360  
EPNYPKSIHS FGFPNFVKKI DAAVENPRFY RTYFFVDNQY WRYDERRQMM DPGYPKLITK 420  
NFQGIGPKID AVFYSKNKYY YFFQGSNQFE YDFLLQRITK TLKSNSWFGC

Seq ID NO: 526 DNA sequence  
Nucleic Acid Accession #: NM\_024423.1  
Coding sequence: 64..2590

1	11	21	31	41	51	
GGCAGGTCTC	GCTCTCGGCA	CCCTCCCGGC	GCCCGCGTTC	TCCTGGCCCT	GCCCGGCATC	60
CCGATGGCCG	CCGCTGGGCC	CCGGCGCTCC	GTGCGCGGAG	CCGTCTGCCT	GCATCTGCTG	120
CTGACCCTCG	TGATCTTCAG	TCGTGATGGT	GAAGCCTGCA	AAAAGGTGAT	ACTTAATGTA	180
CCTTCTAAAC	TAGAGGCAGA	CAAAATAATT	GGCAGAGTTA	ATTTGGAAGA	GTGCTTCAGG	240
TCTGCAGACC	TCATCCGGTC	AAGTGATCCT	GATTTTCAGAG	TTCTAAATGA	TGGGTCACTG	300
TACACAGCCA	GGGCTGTTGC	GCTGTCTGAT	AAGAAAAGAT	CATTTACCAT	ATGGCTTTCT	360
GACAAAAGGA	AACAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAAGAAGGTA	420
TCGAAGACAA	GACACACTAG	AGAAACTGTT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCACCT	480
ATTCCTTGCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTCC	CATTGTTTCT	TCAACAAGTT	540
GAATCTGATG	CAGCACAGAA	CTATACTGTC	TTCTACTCAA	TAAGTGGACG	TGGAGTTGAT	600
AAAGAACCTT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAAATCTATT	TTGCACTCGG	660
CCTGTGGATC	GTGAAGAATA	TGATGTTTTT	GATTTGATTG	CTTATGCGTC	AACTGCAGAT	720
GGATATTCAG	CAGATCTGCC	CCTCCCCTAA	CCCATCAGGG	TAGAGGATGA	AAATGACAAC	780
CACCCTGTTT	TCACAGAAGC	AATTTATAAT	TTTGAAGTTT	TGGAAAAGTAG	TAGACCTGGT	840
ACTACAGTGG	GGGTGGTTTG	TGCCACAGAC	AGAGATGAAC	GGGACACAAT	GCATACGCGC	900
CTGAAATACA	GCATTTTGCA	GCAGACACCA	AGGTCACCTG	GGCTCTTTTC	TGTGCATCCC	960
AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
TCATTGATAA	TGAAAGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTGAT	AGGCACATCA	1080
ACTTGTATCA	TAACAGTAAC	AGATTCAAAT	GATAATGCAC	CCACTTTCAG	ACAAAATGCT	1140
TATGAAGCAT	TTGTAGAGGA	AAATGCATTG	AATGTGAAA	TCTTACGAAT	ACCTATAGAA	1200
GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTTT	AAAGGGAAAT	1260
GAAAATGGAC	ATTTCAAAAT	CAGCACAGAC	AAAGAACTA	ATGAAGGTGT	TCTTTCTGTT	1320
GTAAAGCCAC	TGAATTATGA	AGAAAACCGT	CAAGTGAACC	TGGAAATTGG	AGTAAACAAT	1380
GAAGCGCCAT	TTGCTAGAGA	TATTTCCAGA	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
GTTTCATGTG	GGGATCTGGA	TGAGGGGCCT	GAATGCACTC	CTGCAGCCCA	ATATGTGCGG	1500
ATTAAAGAAA	ACTTAGCAGT	GGGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
AATAGAAATG	GCAATGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
CCCAAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAAATGATA	ATCCACCAGA	AATACTTCAA	1800
GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAGTTTGC	CCAATACTTC	TCCAGAAATC	1920
AGTAGACTGT	GGAGCCTCAC	CAAAGTTTAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
AATGCTGGAT	TTCAAGAATA	TACCATTCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACCTA	GTGTCGTGCG	2100
ACTTCAAGGA	GTACAGGAGT	AATACCTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
ATAGCACTGC	TCTTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTTTGG	TGCAACTAAA	2220
GGGAAACGTT	TTCCTGAAGA	TTTAGCACAG	CAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
CCTGGAGACG	ATAGAGTGTG	CTCTGCCAAT	GATTTTATGA	CCCAAACTAC	CAACAACTCT	2340
AGCTCAAGTT	TTTGTGGTAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAACCATT	2400
GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
ACCCTGGACT	CCTGCAGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACTTACTCG	2520
GAGTGGCACA	GTTTTTACTCA	ACCCCGTCTC	GGTGAAGAAT	CCATTAGAGG	ACACACTGGT	2580
TAAAAATTAA	ACATAAAAGA	AATTGCATCG	ATGTAATCAG	AATGAAGACC	GCATGCCATC	2640
CCAAGATTAT	GTCTTCACTT	ATAACTATGA	GGGAAGAGGA	TCTCCAGCTG	GTTCTGTGGG	2700
CTGCTGCAGT	GAAAAGCAGG	AAGAGATGGG	CCTTGACTTT	TTAAATAAAT	TGGAACCCAA	2760
ATTTATTACA	TTAGCAGAAG	CATGCACAAA	GAGATAATGT	CACAGTGCTA	CAATTAGGTC	2820
TTTGTGAGAC	ATTCTGGAGG	TTTCCAAAAA	TAATATTGTA	AAGTTCAATT	TCAACATGTA	2880
TGTATATGAT	GATTTTTTTC	TCAATTTTGA	ATTATGCTAC	TCACCAATTT	ATATTTTTTA	2940
AGCCAGTTGT	TGCTTATCTT	TTCCAAAAAG	TGAAAAATGT	TAAAACAGAC	AACTGGTAAA	3000
TCTCAAACCT	CAGCACTGGA	ATTAAGGTCT	CTAAAGCATC	TGCTCTTTTT	TTTTTTTACG	3060
GATATTTTAG	TAATAAATAT	GCTGGATAAT	TATTAGTCCA	ACAATAGCTA	AGTTATGCTA	3120
ATATCACATT	ATTATGTATT	CACCTTAAAG	GATAGTTTAA	AAAATAAACA	AGAAATATTG	3180
AGTATCACTA	TGTGAAGAAA	GTTTTGGAAA	AGAAACAATG	AAGACTGAAT	TAAATTAATA	3240
ATGTTGCAGC	TCATAAAGAA	TTGGGACTCA	CCCCTACTGC	ACTACCAAAT	TCATTTGACT	3300
TTGGAGGCAA	AATGTGTTGA	AGTGCCCTAT	GAAGTAGCAA	TTTTCTATAG	GAATATAGTT	3360
GGAAATAAAT	GTGTGTGTGT	ATATTATTAT	TAATCAATGC	AATATTTAAA	ATGAAATGAG	3420
AACAAAGAGG	AAAATGGTAA	AAAGCTGGAA	TGAGGCTGGG	GTATAGTTTG	TCCTACAATA	3480
GAAAAAAGAG	AGAGCTTCCT	AGGCCTGGGC	TCTTAAATGC	TGCATTATAA	CTGAGTCTAT	3540
GAGGAAATAG	TTCCTGTCCA	ATTTGTGTAA	TTTGTTTAAA	ATTGTAAATA	AATTAAACTT	3600
TTCTGGTTTC	TGTGGGAAGG	AAATAGGGAA	TCCAATGGAA	CAGTAGCTTT	GCTTTGAGCT	3660
CTGTTTCAAG	ATTTCTGCAT	CCACAAGTTA	GTAGCAAACT	GGGGAATACT	CGCTGCAGCT	3720
GGGGTTCCCT	GCTTTTTGGT	AGCAAGGGTC	CAGAGATGAG	GTGTTTTTTT	CGGGGAGCTA	3780
ATAACAAAAA	CATTTTAAAA	CTTACCTTTA	CTGAAGTTAA	ATCCTCTATT	GCTGTTTCTA	3840
TTCTCTCTTA	TAGTGACCAA	CATCTTTTAA	ATTTAGATCC	AAATAACCAT	GTCTCTCTAG	3900
AGTTTAGAGG	CTAGAGGGAG	CTGAGGGGAG	GATCTTACTG	AAAGCACCCCT	GGGGAGATTG	3960
ATTGTCCTTA	AACCTAAGCC	CCACAACTTT	GACACCTGAT	CAGGTCTGGG	AGCTACAAAA	4020
TTTCATTTTT	CTCCTCACTG	CCCTTCTTCT	GAGTGGCATT	GGCCTGAATC	AAGGAAAGCC	4080
AGGCCTTGTG	GGCCCCCTTC	TTTCGGCTTT	CTGCTAAAGC	AACACCTCCA	GCAGAGATTG	4140
CCTTAAGTGA	CTCCAGGTTT	TCCACCATCC	TTCAGCGTGA	ATTAATTTTT	AATCAGTTTG	4200
CTTTCTCCAG	AGAAATTTTA	AAATAATAGA	AGAAATAGAA	ATTTTGAATG	TATAAAAGAA	4260
AAAGATCAAG	TTGTCAATTT	AGAACAGAGG	GAACCTTGGG	AGAAAGCAGC	CCAAGTAGGT	4320
TATTTGTACA	GTCAGAGGGC	AACAGGAAGA	TGCAGGCCTT	CAAGGGCAAG	GAGAGGCCAC	4380
AAGGAATATG	GGTGGGAGTA	AAAGCAACAT	CGTCTGCTTC	ATACTTTTTT	CTAGGCTTGG	4440



	CACTGCCTTT	TCCTTTCTCA	GGCCAATGGC	AACTGCCATT	TGAGTCCGGT	GAGGGATCAG	4500
	CCAACCTCTT	CTCTATGGCT	CACCTTATTT	GGAGTGAGAA	ATCAAGGAGA	CAGAGCTGAC	4560
	TGCATGATGA	GTCTGAAGGC	ATTTGCAGGA	TGAGCCTGAA	CTGGTTGTGC	AGAACAAACA	4620
5	AGGCATTTCAT	GGGAATTGTT	GTATTCCTTC	TGCAGCCCTC	CTTCTGGGCA	CTAAGAAGGT	4680
	CTATGAATTA	AATGCCTATC	TAAAATTCTG	ATTTATTCCT	ACATTTTCTG	TTTTCTAATT	4740
	TGACCCTAAA	ATCTATGTGT	TTTAGACTTA	GACTTTTTTAT	TGCCCCCCCC	CCCTTTTTTT	4800
	TTGAGACGGA	GTCTCGCTCT	GACGCACAGG	CTGGAGTGCA	GTGGCTCCGA	TCTCTGCTCA	4860
	CTGAAAGCTC	CGCCTCCCGG	GTTTCATGCC	TTCTCCTGCC	TCAGCCTCCT	GAGTAGCTGG	4920
10	GACTAGAGGC	GCCCCACCAC	ACGCCCCGGT	AATTTTTTGT	ATTTTAAATA	GAGACGGGGT	4980
	TTCACTGTGT	TAGCCAGGAT	GGTCTCGATC	TCCTGACCTC	GTGATCCGCC	TGCCTCGGCC	5040
	TCCCAAAGTG	CTGGGATTAC	AGGCATGACC	CACCGCTCCC	GGCCTTGTTT	TCCGTTTAAA	5100
	GTCGTCTTCT	TTTAATGTAA	TCATTTTGAA	CATGTGTGAA	AGTTGATCAT	ACGAATTGGA	5160
	TCAATCTTGA	AATACTCAAC	CAAAAGACAG	TCGAGAAGCC	AGGGGGAGAA	AGAACTCAGG	5220
	GCACAAAATA	TTGGTCTGAG	AATGGAATTC	TCTGTAAGCC	TAGTTGCTGA	AATTTCTGCT	5280
15	TGTAACAGAG	AGCCAGTTT	ATCTAACGGC	TACTGAAACA	CCCACTGTGT	TTTGCTCACT	5340
	CCCACTCACC	GATCAAAACC	TGCTACCTCC	CCAAGACTTT	ACTAGTGCCG	ATAAACTTTC	5400
	TCAAAGAGCA	ACCAGTATCA	CTTCCTGTG	TATAAAACCT	CTAACCATCT	CTTTGTTCTT	5460
	TGAACATGCT	GAAAACCACC	TGGTCTGCAT	GTATGCCCGA	ATTTGTAATT	CTTTTCTCTC	5520
	AAATGAAAAT	TTAATTTTAG	GGATTCATTT	CTATATTTTC	ACATATGTAG	TATTATTATT	5580
20	TCCTTATATG	TGTAAGGTGA	AATTTATGGT	ATTTGAGTGT	GCAAGAAAAT	ATATTTTAA	5640
	AGCTTTTCATT	TTTCCCCCAG	TGAATGATT	AGAATTTTTT	ATGTAAATAT	ACAGAATGTT	5700
	TTTTCTTACT	TTTATAAGGA	AGCAGCTGTC	TAAAATGCAG	TGGGGTTTGT	TTTGCAATGT	5760
	TTTAAACAGA	GTTTTAGTAT	TGCTATTAAA	AGAAGTTACT	TTGCTTTTAA	AGAAACTTGG	5820
25	CTGCTTAAAA	TAAGCAAAAA	TTGGATGCAT	AAAGTAATAT	TTACAGATGT	GGGGAGATGT	5880
	AATAAAACAA	TATTAACCTG	GCTGCTTAAA	ATAAGCAAAA	ATTGGATGCA	TAAAGTAATA	5940
	TTTACAGATG	TGGGGAGATG	TAATAAAACA	ATATTAACCT	GGTTTCTTGT	TTTTGCTGTA	6000
	TTTAGAGATT	AAATAATTCT	AAGATGATCA	CTTTGCAAAA	TTATGCTTAT	GGCTGGCATG	6060
	GAAATAGAAA	TACTCAATTA	TGTCTTTGTT	GTATTAATGG	GGAATATTTT	GGACAATGTT	6120
30	TCATTATCAA	ATTGTCGACA	TCATTAATAT	ATATTGTAAT	GTTGGGAAGA	GATCACTATT	6180
	TTGAAGCACA	GCTTTACAGA	TGAGTATCTA	TGATACATAT	GTATAATAAA	TTTTGATCGG	6240
	GTATTAAAAA	TATTAGAAGG	TGGTTATAAT	TGCAGAGTAT	TCCATGAATA	GTACACTGAC	6300
	ACAGGGGTTT	TACTTTGAGG	ACCAGTGTAG	TCAAGGGAAA	ACATGAGTTA	AAAAGAAAAG	6360
	CAGGCAATAT	TGCAGTCTTG	ATTCTGCCAC	TTACAGGATA	GATAATGCCT	GAACCTTAAT	6420
35	GACAAGATGA	TCCAACCATA	AAGGTGCTCT	GTGCTTCACA	GTGAATCTTT	TCCCCATGCA	6480
	GGAGTGTGCT	CCCCTACAAA	CGTTAAGACT	GATCATTTCA	AAAATCTATT	AGCTATATCA	6540
	AAAGCCTTAC	ATTTTAATAT	AGGTGGAACC	AAAATTTCAA	TTCCAGTAAC	TTCTATTGTA	6600
	ACCATTATTT	TTGTGTATGT	CTTCAAGAAT	GTTCAATTGA	TTTTTGTTTG	TAATAGTAAA	6660
	ATACCGGATA	CATTTACAGT	GTCCTTCAGT	ATTGATTTGG	TTGAATATTG	GGTCATAATG	6720
40	GTTGAGAAGC	ATGGACACTA	GAGCCAGAA	GCTTGGATAT	GAATCCTGGA	TCTGTCACTT	6780
	ACTTCTGTGT	AGCCTTTGAA	AGGCTCTTAA	TTTCTCTCTC	TAGCTTTCTC	ATTAAATATCA	6840
	ATGAACAATG	CCAGCCTCAT	GGGGTTGTTG	AATGATTAAA	TTAGTTAATA	TACCTAAAGT	6900
	ACATAGAACA	CTGCCTGCAC	ATAGTAAAAG	AATTATAAGT	GTGAGGTAGT	TGGTAAAATT	6960
	ATGTAGTTGG	ATATACTACC	GAACAATATC	TAATCTCTTT	TTAGGGAAAT	AAAGTTTGTG	7020
45	CATATATATA	ATCCCGAAAC	ATG				

Seq ID NO: 527 Protein sequence  
Protein Accession #: NP\_077741.1

	1	11	21	31	41	51	
50							
	MAAAGPRRSV	RGAVCLHLLL	TLVIFSRDGE	ACKKVILNVP	SKLEADKIIG	RVNLEECFRS	60
	ADLIRSSDPD	FRVLNDGSVY	TARAVALSDK	KRSFTIWLSD	KRKQTQKEVT	VLLHQQKKVS	120
	KTRHRTETVL	RRAKRRWAPI	PCSMQENSLG	PFPLFLQQVE	SDAAQNYTVF	YSISGRGVDK	180
55	EPLNLFYIER	DTGNLFCTRP	VDREEDYDVF	LIAYASTADG	YSADLPLPLP	IRVEDENDNH	240
	PVFTEAIYNF	EVLESSRPGT	TVGVVCATDR	DEPDTMHTRL	KYSILQQTPR	SPGLFSVHPS	300
	TGVITTVSHY	LDREVVDKYS	LIMKVQDMDG	QFFGLIGTST	CIITVTDSDN	NAPTFRQNAV	360
	EAFVEENAFN	VEILRIPIED	KDLINTANWR	VNFTILKGNE	NGHFKISTDK	ETNEGVLSV	420
	KPLNYEENRQ	VNLEIGVNNE	APFARDIPRV	TALNRALVTV	HVRDLDEGPE	CTPAAQYVRI	480
60	KENLAVGSKI	NGYKAYDPEN	RNGNGLRYKK	LHDPKGWITI	DEISGSIITS	KILDREVETP	540
	KNELYNITVL	AIDKDDRSCT	GTLAVNIEDV	NDNPPEILQE	YVVICPKPMG	YTDILAVDPD	600
	EPVHGAPFYF	SLPNTSPEIS	RLWSLTKVND	TAARLSYQKN	AGFQEYTIPI	TVKDRAGQAA	660
	TKLLRVNLCE	CTHPTQCRAT	SRSTGVILGK	WAILAILLGI	ALLFSVLLTL	VCGVFGATKG	720
	KRFPEDLAQQ	NLIISNTEAP	GDDRVCASANG	FMTQTNNSS	QGFCGTMSGG	MKNGGQETIE	780
65	MMKGGNQTL	SCRAGHHHT	LDSCRGGHTE	VDNCRYTYSE	WHSFTQPRLG	EESIRGHTG	

Seq ID NO: 528 DNA sequence  
Nucleic Acid Accession #: NM\_001941.2  
Coding sequence: 64..2754

	1	11	21	31	41	51	
70							
	GGCAGGTCTC	GCTCTCGGCA	CCCTCCCGGC	GCCCGCGTTC	TCCTGGCCCT	GCCCGGCATC	60
	CCGATGGCCG	CCGCTGGGCC	CCGCGCTCC	GTGCGCGGAG	CCGTCTGCCT	GCATCTGCTG	120
75	CTGACCCTCG	TGATCTTCAG	TCGTGATGGT	GAAGCCTGCA	AAAAGGTGAT	ACTTAATGTA	180
	CCTTCTAAAC	TAGAGGCAGA	CAAAATAATT	GGCAGAGTTA	ATTTGGAAGA	GTGCTTCAGG	240
	TCTGCAGACC	TCATCCGGTC	AAGTGATCCT	GATTTTCAGAG	TTCTAAATGA	TGGGTCAGTG	300
	TACACAGCCA	GGGCTGTTGC	GCTGTCTGAT	AAGAAAAGAT	CATTTACCAT	ATGGCTTTCT	360
	GACAAAAGGA	AAAAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAAGAAGGTA	420
80	TCGAAGACAA	GACACACTAG	AGAAACTGTT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCACCT	480
	ATTCCTTGCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTCC	CATTGTTTCT	TCAACAAGTT	540
	GAATCTGATG	CAGCACAGAA	CTATACTGTC	TTCTACTCAA	TAAGTGGACG	TGGAGTTGAT	600
	AAAGAACCTT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAAATCTATT	TTGCACTCGG	660
	CCTGTGGATC	GTGAAGAATA	TGATGTTTTT	GATTTGATTT	CTTATGCGTC	AACTGCAGAT	720
85	GGATATTCAG	CAGATCTGCC	CCTCCCCTTA	CCCCTCAGGG	TAGAGGATGA	AAATGACAAC	780
	CACCCTGTTT	TCACAGAAGC	AATTTATAAT	TTTGAAGTTT	TGGAAAGTAG	TAGACCTGGT	840
	ACTACAGTGG	GGGTGGTTTG	TGCCACAGAC	AGAGATGAAC	CGGACACAAT	GCATACGCGC	900
	CTGAAATACA	GCATTTTGCA	GCAGACACCA	AGGTCACCTG	GGCTCTTTTC	TGTGCATCCC	960



	AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
	TCATTGATAA	TGAAAGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTGAT	AGGCACATCA	1080
	ACTTGTATCA	TAACAGTAAC	AGATTCAAAT	GATAATGCAC	CCACTTTCAG	ACAAAATGCT	1140
5	TATGAAGCAT	TTGTAGAGGA	AAATGCATTC	AATGTGGAAA	TCTTACGAAT	ACCTATAGAA	1200
	GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTTT	AAAGGGAAAT	1260
	GAAAATGGAC	ATTTCAAAAT	CAGCACAGAC	AAAGAAACTA	ATGAAGGTGT	TCTTTCTGTT	1320
	GTAAAGCCAC	TGAATTATGA	AGAAAACCGT	CAAGTGAACC	TGGAAATTGG	AGTAAACAAT	1380
	GAAGCGCCAT	TTGCTAGAGA	TATTTCCAGA	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
10	GTTTCATGTGA	GGGATCTGGA	TGAGGGGCCT	GAATGCACTC	CTGCAGCCCA	ATATGTGCGG	1500
	ATTAAAGAAA	ACTTAGCAGT	GGGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
	AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCAAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAAATGATA	ATCCACCAGA	AATACTTCAA	1800
15	GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAGTTTGC	CCAATACTTC	TCCAGAAATC	1920
	AGTAGACTGT	GGAGCCTCAC	CAAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGAATA	TACCATTCCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
20	GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACTCA	GTGTCGTGCG	2100
	ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCTTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTGTGG	TGCAACTAAA	2220
	GGGAAACGTT	TTCTGAAGA	TTTAGCAGAG	CAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
	CCTGGAGACG	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAAACTAC	CAACAACCTCT	2340
25	AGCCAAGGTT	TTTGTGTGAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAACCATT	2400
	GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCCTGGACT	CCTGCAGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACTTACTCG	2520
	GAGTGGCACA	GTTTTACTCA	ACCCCGTCTC	GGTGAAAAAT	TGCATCGATG	TAATCAGAAT	2580
	GAAGACCGCA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGATCT	2640
30	CCAGCTGGTT	CTGTGGGCTG	CTGCAGTGAA	AAGCAGGAAG	AAGATGGCCT	TGACTTTTTTA	2700
	AATAATTTGG	AACCCAAAT	TATTACATTA	GCAGAAGCAT	GCACAAAGAG	ATAATGTCAC	2760
	AGTGCTACAA	TTAGGTCCTT	GTCAGACATT	CTGGAGGTTT	CCAAAAATAA	TATTGTAAAG	2820
	TTCAATTTCA	ACATGTATGT	ATATGATGAT	TTTTTTCTCA	ATTTTGAATT	ATGCTACTCA	2880
	CCAATTTATA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAAATGTTAA	2940
35	AACAGACAAC	TGGTAAATCT	CAAACTCCAG	CAGTGGAATT	AAGGTCTCTA	AAGCATCTGC	3000
	TCTTTTTTTT	TTTTACGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAACA	3060
	ATAGCTAAGT	TATGCTAATA	TCACATTATT	ATGTATTACAC	TTTAAGTGAT	AGTTTAAAAA	3120
	ATAAACAAGA	AATATTGAGT	ATCACTATGT	GAAGAAAGTT	TTGGAAAAGA	AACAATGAAG	3180
	ACTGAATTAA	ATTAAAAATG	TTGCAGCTCA	TAAAGAATTG	GGACTCACCC	CTACTGCAC	3240
40	ACCAAATTCA	TTTGACTTTG	GAGGCAAAAT	GTGTTGAAGT	GCCCTATGAA	GTAGCAATTT	3300
	TCTATTGGAA	TATAGTTGGA	AATAAATGTG	TGTGTGTATA	TTATTTATTA	TCAATGCAAT	3360
	ATTTAAAATG	AAATGAGAAC	AAAGAGGAAA	ATGGTAAAAA	CTTGAAATGA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCCTAGG	CCTGGGCTCT	TAAATGCTGC	3480
	ATTATAACTG	AGTCTATGAG	GAAATAGTTC	CTGTCCAATT	TGTGTAATTT	GTTTAAATTT	3540
45	GTAAATAAAT	TAAACTTTTC	TGGTTTCTGT	GGGAAGGAAA	TAGGGAAATCC	AATGGAACAG	3600
	TAGCTTTGCT	TTGCAGTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATACTCGC	TGCAGCTGGG	GTTCCCTGCT	TTTTGGTAGC	AAGGGTCCAG	AGATGAGGTG	3720
	TTTTTTTTCG	GGAGCTAATA	ACAAAAACAT	TTTTAAACTT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATTGCT	GTTTCTATT	TCTCTTATAG	TGACCAACAT	CTTTTAAATT	TAGATCCAAA	3840
50	TAACCATGTC	CTCCTAGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
	GCACCCTGGG	GAGATTGATT	GTCCTTAAAC	CTAAGCCCCA	CAAACCTGAC	ACCTGATCAG	3960
	GTCTGGGAGC	TACAAAATTT	CATTTTCTC	CTCACTGCCC	TTCTTCTGAG	TGGCATTGGC	4020
	CTGAATCAAG	GAAAGCCAGG	CCTTGTGGGG	CCCCTTCTTT	CGGCTTCTG	CTAAAGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAAGTGACTC	CAGGTTTTC	ACCATCCTTC	AGCGTGAATT	4140
55	AATTTTAAAT	CAGTTTGCTT	TCTCCAGAGA	AATTTTAAAA	TAATAGAAGA	AATAGAAATT	4200
	TTGAATGTAT	AAAAGAAAAA	GATCAAGTTG	TCATTTTAGA	ACAGAGGGAA	CTTTGGGAGA	4260
	AAGCAGCCCA	AGTAGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCTTCAA	4320
	GGGCAAGGAG	AGGCCACAAG	GAATATGGGT	GGGAGTAAAA	GCAACATCGT	CTGCTTCATA	4380
	CTTTTTCCTA	GGCTTGGCAC	TGCCTTTTCC	TTTCTCAGGC	CAATGGCAAC	TGCCATTTGA	4440
60	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGCTCAC	CTTATTTGGA	GTGAGAAATC	4500
	AAGGAGACAG	AGCTGACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACTG	4560
	GTTGTGCAGA	ACAAACAAGG	CATTATGAGG	AATTGTTGTA	TTCTTCTGCT	AGCCCTCCTT	4620
	CTGGGCACTA	AGAAGGTCTA	TGAATTAAAT	GCCTATCTAA	AATTCTGATT	TATTCCTACA	4680
	TTTTCTGTTT	TCTAATTTGA	CCCTAAAAAT	TATGTGTTTT	AGACTTAGAC	TTTTTATTGC	4740
65	CCCCCCCCC	TTTTTTTTTG	AGACGGAGT	TCGCTCTGAC	GCACAGGCTG	GAGTGCAGTG	4800
	GCTCCGATCT	CTGCTCACTG	AAAGCTCCGC	CTCCCGGTT	CATGCCATTC	TCCTGCCTCA	4860
	GCCTCCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCACG	CCCGGCTAAT	TTTTTGTATT	4920
	TTTAATAGAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTCGATCTCC	TGACCTCGTG	4980
	ATCCGCTGTC	CTCGGCCTCC	CAAAGTGCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040
70	CTTGTTTTCC	GTTTAAAGTC	GTCCTTCTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAAGT	5100
	TGATCATACG	AATTTGGATCA	ATCTTGAAAT	ACTCAACCAA	AAGACAGTCG	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGGCA	CAAAATATTG	GTCTGAGAAT	GGAATTTCTT	GTAAGCCTAG	5220
	TTGCTGAAAT	TTCTTGCTGT	AACCAGAAAG	CAGTTTTATC	TAACGGCTAC	TGAAACACCC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACTCACCG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
75	CTAGTGCCGA	TAACTTTCTT	CAAAGAGCAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
	TAACCATCTC	TTTGTCTCTT	GAACATGCTG	AAAACCACCT	GGTCTGCATG	TATGCCCGAA	5460
	TTTGTAATTC	TTTCTCTCTA	AATGAAAATT	TAATTTTAGG	GATTCATTTT	TATATTTTCA	5520
	CATATGTAGT	ATTATTATTT	CCTTATATGT	GTAAGGTGAA	ATTTATGGTA	TTTGAGTGTG	5580
	CAAGAAAAATA	TATTTTAAAA	GCTTTTCAAT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640
80	TGTAAATATA	CAGAATGTTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAAATGCAGT	5700
	GGGGTTTGTT	TTGCAATGTT	TTAAACAGAG	TTTTAGTATT	GCTATTAAAA	GAAGTTACTT	5760
	TGCTTTTTAAA	GAAACTTGGC	TGCTTAAAAAT	AAGCAAAAAT	TGGATGCATA	AAGTAATATT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAACAAT	ATTAACCTGG	TTTCTTGTTT	TTGCTGTATT	5880
	TAGAGATTAA	ATAATTCTAA	GATGATCACT	TTGCAAAAAT	ATGCTTATGG	CTGGCATGGA	5940
85	AATAGAAATA	CTCAATTATG	TCTTTGTGTT	ATTAATGGGG	AATATTTTGG	ACAATGTTTC	6000
	ATTATCAAAAT	TGTCGACATC	ATTGTAATGT	TGGGAAGAGA	TCACTATTTT		6060
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAATT	TTGATCGGGT	6120
	ATTAAAAGTA	TTAGAAGGTG	GTTATAATTG	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180



AGGGGTTTTTA CTTTGAGGAC CAGTGTAGTC AAGGGAAAAC ATGAGTTAAA AAGAAAAGCA 6240  
GGCAATATTG CAGTCTTGAT TCTGCCACTT ACAGGATAGA TAATGCCTGA ACTTTAATGA 6300  
CAAGATGATC CAACCATAAA GGTGCTCTGT GCTTCACAGT GAATCTTTTC CCCATGCAGG 6360  
5 AGTGTGCTCC CCTACAAACG TTAAGACTGA TCATTTCAAA AATCTATTAG CTATATCAAA 6420  
AGCCTTACAT TTTAATATAG GTTGAACCAA AATTTCAATT CCAGTAACTT CTATTGTAAC 6480  
CATTATTTTT GTGTATGTCT TCAAGAATGT TCATTGGATT TTTGTTTGTG ATAGTAAAT 6540  
ACCGGATACA TTTCACGTGT CCTTCAGTAT TGATTGGTT GAATATGGG TCATAATGGT 6600  
TGAGAAGCAT GGACACTAGA GCCAGAATGC TTGGATATGA ATCCTGGATC TGTCACTTAC 6660  
10 TTCTGTGTGA CCTTTGAAAG GCTACTTATT TCCTCTCTTA GCTTTCTCAT TAAAATCAAT 6720  
GAACAATGCC AGCCTCATGG GGTGTTGAA TGATTAAATT AGTTAATATA CCTAAAGTAC 6780  
ATAGAACACT GCCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840  
GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTTT AGGGAAATAA AGTTTGTGCA 6900  
TATATATAAT CCCGAAACAT G

Seq ID NO: 529 Protein sequence  
Protein Accession #: NP\_001932.1

1 11 21 31 41 51  
| | | | | |  
20 MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60  
ADLIRSSDPD FRVLNDGSVY TARAVALS DK KRSFTIWLSD KRKQTQKEVT VLEHQKKVS 120  
KTRHRTRETVL RRAKRRWAPI PCSMQENSLG PFPLFLQQVE SDAAQNYTVF YSISGRGV DK 180  
EPLNLFYIER DTGNLFCTRP VDREEDYDVF LIAVASTADG YSADLPLPLP IRVEDENDNH 240  
25 PVFTEAIYNF EVLESSRPGT TVGVVCATDR DEPDTMHTRL KYSILQQTTP SPGLFSVHPS 300  
TGVITTVSHY LDREVVDKYS LIMKVQDMDG QFFGLIGTST CIITVDSND NAPTFRQNAY 360  
EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKGNE NGHFKISTDK ETNEGVL SVV 420  
KPLNYEENRQ VNLEIGVNE APFARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI 480  
KENLAVGSKI NGYKAYDPEN RENGNGLYKK LHDPKGWITI DEISGSIITS KILDREVETP 540  
KNELYNITVL AIDKDDRSCT GTLAVNIEDV NDNPP EILQE YVVICPKMG YTDILAVDPD 600  
30 EPVHGAPFYF SLPNTSPEIS RLWSLT KVND TAARLSYQKN AGFQEYTIPI TVKDRAGQAA 660  
TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLTL VCGVFGATKG 720  
KRPEDLAQQ NLIISNTEAP GDDRVC SANG FMTQTTNNS QGFCGTMGSG MKNGGQETIE 780  
MMKGGNQTL SCRAGHHHT LDSCRGGHTE VDNCRYTYSE WHSFTQPRLG EKLHRCNQNE 840  
35 DRMP SQDYVL TYNYEGRGSP AGSVGCCSEK QEEDGLDFLN NLEPKFITLA EACTKR

Seq ID NO: 530 DNA sequence  
Nucleic Acid Accession #: NM\_016583.2  
Coding sequence: 72..842

1 11 21 31 41 51  
| | | | | |  
45 GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60  
TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCCAGA 120  
CCATGGCCCA GTTTGGAGGC CTGCCCGTGC CCCTGGACCA GACCCTGCC TTGAATGTGA 180  
ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GCCCTCAGCA 240  
ATGGCCTGCT GTCTGGGGGC CTGTTGGGCA TTCTGGAAAA CCTTCCGCTC CTGGACATCC 300  
TGAAGCCTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGGA AAAGTGACGT 360  
50 CAGTGATTCC TGGCCTGAAC AACATCATTG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420  
AACTTGGCCT TGTGCAGAGC CCTGATGGCC ACCGTCTCTA TGTACCATC CCTCTCGGCA 480  
TAAAGCTCCA AGTGAATACG CCCCTGGTCG GTGCAAGTCT GTTGAGGCTG GCTGTGAAGC 540  
TGGACATCAC TGCAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC 600  
TTGGTGACTG CACCCATTCC CCTGGAAGCC TGCAAATTTT TCTGCTTGAT GGA CTGGCC 660  
55 CCCTCCCAT TCAAGGTCTT CTGGACAGCC TCACAGGGAT CTTGAATAAA GTCCTGCCTG 720  
AGTTGGTTCA GGGCAACGTG TGCCCTCTGG TCAATGAGGT TCTCAGAGGC TTGGACATCA 780  
CCCTGGTGCA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTC ATCAAGGTCT 840  
AAGCCTTCCA GGAAGGGGCT GGCTCTGCT GAGCTGCTT CCAGTGCTCA CAGATGGCTG 900  
GCCCATGTGC TGAAGATGA CACAGTTGCC TTCTCTCCGA GGAACCTGCC CCCTCTCCTT 960  
60 TCCACCAGG CGTGTGTAAC ATCCCATGTG CCTCACCTAA TAAAATGGCT CTTCTTCTGC 1020  
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 531 Protein sequence  
Protein Accession #: NP\_057667.1

1 11 21 31 41 51  
| | | | | |  
65 MFQTGGLIVF YGLLAQTMAQ FGGLPVPLDQ TLPLNVN PAL PLSPTGLAGS LTNALSNGLL 60  
SGGLLGILEN LPLLDILKPG GGTSGGLLG LGKVT SVIP GLNNIIDIKV TDPQLLELGL 120  
70 VQSPDGHRLY VTIP LGIKLQ VNTPLVGASL LRLAVKLDIT AEILAVRDKQ ERIHLVLGDC 180  
THSPGSLQIS LLDGLGPLPI QGLDLSLTGI LNKVLP ELVQ GNVCLPVNEV LRGLDITLVH 240  
DIVNMLIHGL QFVIKV

Seq ID NO: 532 DNA sequence  
Nucleic Acid Accession #: NM\_004363.1  
Coding sequence: 115..2223

1 11 21 31 41 51  
| | | | | |  
80 CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 60  
TCCTGGAAC TCAAGCTCTT TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 120  
TCTCCCTCGG CCCCTCCCA CAGATGGTGC ATCCCCTGGC AGAGGCTCCT GCTCACAGCC 180  
TCACTTCTAA CCTTCTGGAA CCCGCCACC ACTGCCAAGC TCACTATTGA ATCCACGCCG 240  
85 TTCAATGTCG CAGAGGGGAA GGAGGTGCTT CTACTTGTCC ACAATCTGCC CCAGCATCTT 300  
TTTGGCTACA GCTGGTACAA AGGTGAAAGA GTGGATGGCA ACCGTCAAAT TATAGGATAT 360  
GTAATAGGAA CTCAACAAGC TACCCAGGG CCCGCATACA GTGGTCGAGA GATAATATAC 420  
CCCAATGCAT CCCTGCTGAT CCAGAACATC ATCCAGAATG ACACAGGATT CTACACCCTA 480



5 CACGTCATAA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCCG GGTATACCCG 540  
GAGCTGCCCC AGCCCTCCAT CTCCAGCAAC AACTCCAAAC CCGTGGAGGA CAAGGATGCT 600  
GTGGCCTTCA CCGTGTGAACC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAAT 660  
CAGAGCCTCC CGGTCACTCC CAGGCTGCAG CTGTCCAATG GCAACAGGAC CCTCACTCTA 720  
TTCAATGTCA CAAGAAATGA CACAGCAAGC TACAAATGTG AAACCCAGAA CCCAGTGAGT 780  
GCCAGGCGCA GTGATTCAGT CATCCTGAAT GTCCTCTATG GCCCGGATGC CCCCACCAT 840  
TCCCCTCTAA ACACATCTTA CAGATCAGGG GAAAATCTGA ACCTCTCCTG CCACGCAGCC 900  
TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTTCCAGCA ATCCACCCAA 960  
10 GAGCTCTTTA TCCCCAACAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT 1020  
AACTCAGACA CTGGCCTCAA TAGGACCACA GTCACGACGA TCACAGTCTA TGCAGAGCCA 1080  
CCCAAACCCCT TCATCACCAG CAACAACCTCC AACCCCGTGG AGGATGAGGA TGCTGTAGCC 1140  
TTAACCTGTG AACCTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC 1200  
CTCCCGGTCA GTCCAGGCT GCAGCTGTCC AATGACAACA GGACCCCTCAC TCTACTCAGT 1260  
15 GTCACAAGGA ATGATGTAGG ACCCTATGAG TGTGGAATCC AGAACGAATT AAGTGTGAC 1320  
CACAGCGACC CAGTCATCCT GAATGCTCTC TATGGCCCAG ACGACCCAC CATTTCCTCC 1380  
TCATACACCT ATTACCGTCC AGGGGTGAAC CTCAGCCTCT CCTGCCATGC AGCCTCTAAC 1440  
CCACCTGCAC AGTATTCTTG GCTGATTGAT GGGAACTATC AGCAACACAC ACAAGAGCTC 1500  
TTTATCTCCA ACATCACTGA GAAGAACAGC GGACTCTATA CCTGCCAGGC CAATAACTCA 1560  
20 GCCAGTGGCC ACAGCAGGAC TACAGTCAAG ACAATCACAG TCTCTGCGGA GCTGCCCAAG 1620  
CCCTCCATCT CCAGCAACAA CTCCAAACCT GTGGAGGACA AGGATGCTGT GGCCTTCACC 1680  
TGTGAACCTG AGGCTCAGAA CACAACCTAC CTGTGGTGGG TAAATGGTCA GAGCCTCCCA 1740  
GTCAGTCCCA GGCTGCAGCT GTCCAATGGC AACAGGACCC TCACTCTATT CAATGTCACA 1800  
AGAAATGACG CAAGAGCCTA TGTATGTGGA ATCCAGAACT CAGTGAGTGC AAACCGCAGT 1860  
25 GACCCAGTCA CCCTGGATGT CCTCTATGGG CCGGACACCC CCATCATTTT CCCCCCAGAC 1920  
TCGTCTTACC TTTGGGAGC GAACCTCAAC CTCTCCTGCC ACTCGGCCTC TAACCCATCC 1980  
CCGAGTATT CTTGGCGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTTATC 2040  
GCCAAAATCA CGCCAAATAA TAACGGGACC TATGCCTGTT TTGTCTCTAA CTTGGCTACT 2100  
GGCCGCAATA ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TTCTCCTGGT 2160  
30 CTCTCAGCTG GGGCCACTGT CCGCATCATG ATTGGAGTGC TGGTTGGGGT TGCTCTGATA 2220  
TAGCAGCCCT GGTGTAGTTT CTTCAATTTCA GGAAGACTGA CAGTTGTTTT GCTTCTTCCT 2280  
TAAAGCATTT GCAACAGCTA CAGTCTAAAA TTGCTTCTTT ACCAAGGATA TTTACAGAAA 2340  
AGACTCTGAC CAGAGATCGA GACCATCCTA GCCAACATCG TGAAACCCCA TCTCTACTAA 2400  
AAATACAAAA ATGAGCTGGG CTTGGTGGCG CGCACCTGTA GTCCAGTTA CTCGGGAGGC 2460  
35 TGAGGCAGGA GAATCGCTTG AACCCGGGAG GTGGAGATTG CAGTGAGCCC AGATCGCACC 2520  
ACTGCACTCC AGTCTGGCAA CAGAGCAAGA CTCCATCTCA AAAAGAAAAG AAAAGAAGAC 2580  
TCTGACCTGT ACTCTGAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTTCCAA 2640  
AACTTTAATG AACTAATGA CAGCTTCATG AAAGTGTCCA CCAAGATCAA GCAGAGAAAA 2700  
TAATTAATTT CATGGGACTA AATGAATAA TGAGGATTGC TGATTCTTTA AATGTCTTGT 2760  
40 TTCCAGATT TCAGGAACT TTTTCTCTT TAAGCTATCC ACTCTTACAG CAATTTGATA 2820  
AAATATACTT TTGTGAACAA AATATGAGC ATTTACATT TCTCCCTATG TGGTCGCTCC 2880  
AGACTTGGGA AACTATTCAT GAATATTTAT ATTGTATGGT AATATAGTTA TTGCACAAGT 2940  
TCAATAAAAA TCTGCTCTTT GTATAACAGA AAAA

Seq ID NO: 533 Protein sequence  
Protein Accession #: NP\_004354.1

50 1 11 21 31 41 51  
| | | | | |  
MESPSAPPHR WCIWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE VLLLVHNLPO 60  
HLFGYSWYKG ERVDGNRQII GYVIGTQQAT PGPAYSGREI IYPNASLLIQ NIIQNDTGFI 120  
TLHVIKSDLV NEEATGQFRV YPELPKPSIS SNNSKPVEDK DAVAFCEPE TQDATYLVWV 180  
NNQSLPVSFR LQLSNGNRTL TLFNVTRNDT ASYKCTQNP VSARRSDSVI LNVLYGPDAP 240  
TISPLNTSYR SGENLNLSCH AASNPPAQS WFNKGTQQS TQELFIPNIT VNNSGSYTCQ 300  
55 AHNSTGLNR TTVTTITVYA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ NTTYLVWVNN 360  
QSLPVSFRLQ LSNDNRTLTL LSVTRNDVGP YECGIQNELS VDHSDPVILN VLYGPDPTI 420  
SPSYTYRPG VNLSLSCHAA SNPPAQYSWL IDGNIQHTQ ELFI SNITEK NSGLYTCQAN 480  
NSASGHSRTT VKTITVSSEL PKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLVWVNGQS 540  
LPVSPRLQLS NGNRTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDPTIISP 600  
60 PDSSYLSGAN LNLCHSASN PSPQYSWRIN GIPQHTQVL FIAKITPNNN GTYACFVSNL 660  
ATGRNNSIVK SITVSASGTS PGLSAGATVG IMIGVLVGVA LI

Seq ID NO: 534 DNA sequence  
Nucleic Acid Accession #: NM\_006952.1  
Coding sequence: 11..793

70 1 11 21 31 41 51  
| | | | | |  
AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTTTCGTTGC TTCCAGGGCC TGCTGATTTT 60  
TGGAATATGT ATTATTGGTT GTTGCAGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120  
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG 180  
GGCTGCCTGG ATCGGCATAT TTGTGGGCAT CTGCTCTTTC TGCTGTCTG TTTAGGCAT 240  
TGATAGCATC ATGAAGTCCA GCAGGAAAAA TCTTCTGGCG TATTTTATTC TGATGTTTAT 300  
AGTATATGCC TTTGAAGTGG CATCTTGTAT CACAGCAGCA ACACAACGAG ACTTTTTTAC 360  
75 ACCCAACCTC TTCCTGAAGC AGATGCTAGA GAGGTACCAA AACAACAGCC CTCCAAACAA 420  
TGATGACCAG TGGAATAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480  
CAATTGCTGT GGCCTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CCTTCCGGAC 540  
TGAGAATAAT GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA 600  
AGAACCTCTC AACCTGGAGG CTTGTAAACT AGGCGTGCCT GGTTTTTTATC ACAATCAGGG 660  
80 CTGCTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CCTGGTTTGG 720  
ATTTGCCATT CTCTGCTGGA CTTTCTGGGT TCTCCTGGGT ACCATGTTCT ACTGGAGCAG 780  
AATTGAATAT TAAGAA

Seq ID NO: 535 Protein sequence  
Protein Accession #: NP\_008883.1

85 1 11 21 31 41 51  
| | | | | |



MAKDNSTVRC FQGLLIFGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAW 60  
 IGIFVIGICLF CLSVLGIVGI MKSSRKILLA YFILMFIVYA FEVASCITAA TQRDFFTPNL 120  
 FLKQMLERYQ NNSPPNDDQ WKNNGVTKTW DRLMLQDNCC GVNGPSDWQK YTSAFRTENN 180  
 DADYPWPRQC CVMNMLKEPL NLEACKLGVP GFYHNQGCYE LISGPMNRHA WGVAVFGFAI 240  
 LCWTFWVLLG TMFYWSRIEY

Seq ID NO: 536 DNA sequence  
 Nucleic Acid Accession #: NM\_002638.1  
 Coding sequence: 120..473

1 11 21 31 41 51  
 | | | | |  
 CAATACAGCT AAGGAATTAT CCCTTGTAAT TACCACAGAC CCGCCCTGGA GCCAGGCCAA 60  
 GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120  
 TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGCTGGGACG CTGGTTCTAG 180  
 AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240  
 TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAAGTAA AGGTCAAGAT AAAGTCAAAG 300  
 CGCAAGAGCC AGTCAAAGGT CCAGTCTCCA CTAAGCCTGG CTCCTGCCCC ATTATCTTGA 360  
 TCCGGTGC GC CATGTTGAAT CCCCCTAAC GCTGCTTGAA AGATACTGAC TGCCAGGAA 420  
 TCAAGAAGTG CTGTGAAGGC TCTTGCGGGA TGGCCTGTTT CGTTCCCCAG TGAAGGGAGC 480  
 CGGTCTTTGC TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540  
 TGCTGCCCTT CCCCTTCCCA CACTGTCCAT TCTTCTCTCC ATTACAGGATG CCCACGGCTG 600  
 GAGCTGCCTC TCTCATCCAC TTTCCAATAA A

Seq ID NO: 537 Protein sequence  
 Protein Accession #: NP\_002629.1

1 11 21 31 41 51  
 | | | | |  
 MRASSFLIVV VFLIAGTLVL EAAVTGVPVK GQDTVKGRVP FNGQDPVKGQ VSVKGQDKVK 60  
 AQEPVKGPVS TKPGSCPIIL IRCAMLNPPN RCLKDTDCPG IKKCEGSCG MACFVPQ

Seq ID NO: 538 DNA sequence  
 Nucleic Acid Accession #: NM\_001793.2  
 Coding sequence: 71..2560

1 11 21 31 41 51  
 | | | | |  
 AAAGGGGCAA GAGCTGAGCG GAACACCGGC CCGCCGTCGC GGCAGCTGCT TCACCCCTCT 60  
 CTCTGCAGCC ATGGGGCTCC CTCGTGGACC TCTCGCGTCT CTCCTCCTTC TCCAGGTTTG 120  
 CTGGCTGCAG TGCGCGGCCT CCGAGCCGTG CCGGGCGGTC TTCAGGGAGG CTGAAGTGAC 180  
 CTTGGAGGCG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAAG TATTCATGGG 240  
 CTGCCCTGGG CAAGAGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300  
 TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360  
 ATCCAAACGT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420  
 TGAAAATGGC AAGGTCCCTT TCCCCCAGAG ACTGAATCAG CTCAGTCTA ATAAAGATAG 480  
 AGACACCAAG ATTTTCTACA GCATCACGGG GCGGGGGGCA GACAGCCCCC CTGAGGGTGT 540  
 CTTGCTGTGA GAGAAGGAGA CAGGCTGGTT GTTGTGTAAT AAGCCACTGG ACCGGGAGGA 600  
 GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTGAGAG AATGGTGCCCT CAGTGGAGGA 660  
 CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720  
 GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCCTACCA GGTACTTCTG TGATGCAGGT 780  
 GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAT GGGGTGGTTG CTTACTCCAT 840  
 CCATAGCCAA GAACCAAGG ACCCACACGA CCTCATGTTT ACCATTCACC GGAGCACAGG 900  
 CACCATCAGC GTCATCTCCA GTGGCCTGGA CCGGGAAAAA GTCCCTGAGT ACACACTGAC 960  
 CATCCAGGCC ACAGACATGG ATGGGGACGG CTCCACCACC ACGGCAGTGG CAGTAGTGGA 1020  
 GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080  
 GCCTGAGAAT GCAGTGGGCC ATGAGGTGCA GAGGCTGACG GTCAGTGATC TGGACGCCCC 1140  
 CAACCTACCA GCGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACG GGGACCATTT 1200  
 TACCATCACC ACCCAACCTG AGAGCAACCA GGGCATCCTG ACAACCAGGA AGGGTTTGGG 1260  
 TTTTGAGGCC AAAAACCAGC ACACCCTGTA CGTTGAAGTG ACCAACGAGG CCCCTTTTGT 1320  
 GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAC GTGGAGGATG TGAATGAGGC 1380  
 ACCTGTGTTT GTCCACCCCT CCAAAGTCGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440  
 GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCAGCTACCG 1500  
 CATCTTGAGA GACCCAGCAG GGTGGCTAGC CATGTACCCA GACAGTGGGC AGGTACAGC 1560  
 TGTGGGCACC CTGACCGTG AGGTAGGACA GTTTGTGAGG AACAACTCT ATGAAGTCAT 1620  
 GGTCTTGGCC ATGGACAATG AAGGCCCTCC CACCACTGGC ACGGGAACCC TTCTGCTAAC 1680  
 ACTGATTGAT GTCAATGACC ATGGCCAGT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740  
 CCAAAGCCCT GTGCGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800  
 CCCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860  
 GGAAGGTGAC ACAGTGGTCT TGTCCCTGAA GAAGTTCCTG AAGCAGGATA CATATGACGT 1920  
 GCACCTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980  
 GTGCGACTGC CATGGCCATG TCGAAACCTG CCCTGGACCC TGGAAGGGAG GTTTTCATCCT 2040  
 CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCTCTCTG CTGGTGCTGC TTTTGTGTTG 2100  
 GAGAAAGAAG CGGAAGATCA AGGAGCCCTT CCTACTCCCA GAAGATGACA CCCGTGACAA 2160  
 CGTCTTCTAC TATGGCGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220  
 GCTCCACCGA GGTCTGGAGG CCAGGCCGGA GGTGGTTCTC CGCAATGACG TGGCACCAAC 2280  
 CATCATCCCG ACACCATGT ACCGTCTCTG CCGAGCCAAC CCAGATGAAA TCGGCAACTT 2340  
 TATAATTGAG AACCTGAAG CGGCTAACAC AGACCCACA GCCCCGCCCT ACACACCTC 2400  
 CTTGGTGTTC GACTATGAGG GCAGCGGCTC CGACGCGCGG TCCCTGAGCT CCCTCACCTC 2460  
 CTCCGCTCC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCCGCTTCAA 2520  
 GAAGCTGGCA GACATGTACG GTGGCGGGGA GGACGACTAG GCGGCCTGCC TGCAGGGCTG 2580  
 GGGACCAAC GTCAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG 2640  
 GACTTCGGAG CTTGTCAGGA AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700  
 ACGTTAGAGT GGTGCTTCC TTAGCTTTT AGGATGGAGG AATGTGGGCA GTTTGACTTC 2760  
 AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCCTCAG AGGCCAAGTT TCCAGAAGCC 2820  
 TCTTACCTGC CGTAAAATGC TCAACCCTGT GTCTGGGCC TGGGCTGCT GTGACTGACC 2880  
 TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTTT 2940



TTTTTTTAAAT GCTATCTTCA AAACGTTAGA GAAAGTTCTT CAAAAGTGCA GCCCAGAGCT 3000  
 GCTGGGCCCA CTGGCCGTCC TGCATTTCTG GTTTCAGAC CCAATGCCT CCCATTCGGA 3060  
 TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGGTGGCC CCTTATTTT TATTTTCCCT 3120  
 GTTGCCTTGC TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTTTAT 3180  
 TAAAGAACT TTTCCAGAA AAAAA

Seq ID NO: 539 Protein sequence  
 Protein Accession #: NP\_001784.2

1 11 21 31 41 51  
 MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG 60  
 QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120  
 KGPFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLLN KPLDREEIAK 180  
 YELFGHAVSE NGASVEDPMN ISIIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240  
 DEDDAIYTYN GVVAYSISHSQ EPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300  
 TDMGDGSGST TAVAVVEILD ANDNAPMFDP QKYEAHVPEN AVGHEVQRLT VTDLDAPNSP 360  
 AWRATYILIM GDDGDHFTIT THPESNQIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420  
 PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR 480  
 DPAGWLAMPD DSGQVTAVGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLLTID 540  
 VNDHGPVPEP RQITICNQSP VRQVLNITDK DLSPTSPFQ AQLTDDSDIY WTAEVNEEGD 600  
 TVVLSLKFL KQDTYDVHLS LSDHGNKEQL TVIRATVCDC HGHVETCPGP WKGGFILPVL 660  
 GAVLALLFLL LVLLLLVRKK RKIKEPLLLP EDDTRDNVYF YGEEGGGEED QDYDITQLHR 720  
 GLEARPEVVL RNDVAPTII TPMPYRPRAN PDEIGNFIE NLKAANTDPT APPYDTLLVF 780  
 DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

Seq ID NO: 540 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..672

1 11 21 31 41 51  
 ATGAGGCTCC AAAGACCCCG ACAGGCCCCG CGCGGTGGGA GGCGCGCGCC CCGGGCGGG 60  
 CGGGGCTCCC CCTACCGGCC AGACCCGGGG AGAGGCGCGC GGAGGCTGCG AAGGTTCCAG 120  
 AAGGGCGGGG AGGGGCGGCC GCGCGCTGAC CCTCCCTGGG CACCGCTGGG GACGATGGCG 180  
 CTGCTCGCCT TGCTGCTGGT CGTGGCCCTA CCGCGGTGT GGACAGACGC CAACCTGACT 240  
 GCGAGACAAC GAGATCCAGA GGAATCCCAG CGAACGGACG AGGGTGACAA TAGAGTGTGG 300  
 TGTCATGTTT GTGAGAGAGA AAACACTTTC GAGTGCCAGA ACCCAAGGAG GTGCAAATGG 360  
 ACAGAGCCAT ACTGCGTTAT AGCGGCCGTG AAAATATTTC CACGTTTTTT CATGTTTGGC 420  
 AAGCAGTGCT CCGCTGGTTG TGCAGCATG TGCAGACCCA AGCCAGAGGA GAAGCGGTTT 480  
 CTCCTGGAAG AGCCCATGCC CTCTTTTAC CTCAAGTGT GTAAAATTCG CTAAGTCAAT 540  
 TTAGAGGGGC CACCTATCAA CTCATCAGTG TTCAAAGAAT ATGCTGGGAG CATGGGTGAG 600  
 AGCTGTGGTG GGCTGTGGCT GGCCATCCTC CTGCTGCTGG CCTCCATTGC AGCCGGCCTC 660  
 AGCCTGTCTT GA

Seq ID NO: 541 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MRLQRPRQAP AGRRAPRGG RGSYPYRDPG RGARRLRRFQ KGGEGAPRAD PPWAPLGTMA 60  
 LLALLLVVAL PRVWTDANLT ARQRDPEDSQ RTDEGDNRVW CHVCERENTF ECQNPRRCKW 120  
 TEPYCVIAAV KIFPRPFMVA KQCSAGCAAM ERPKPEEKRF LLEEMPFFY LKCKIRYCN 180  
 LEGPPINSSV FKEYAGSMGE SCGGLWLAIL LLLASIAAGL SLS

Seq ID NO: 542 DNA sequence  
 Nucleic Acid Accession #: XM\_035292.2  
 Coding sequence: 53..1576

1 11 21 31 41 51  
 GCTCGCTGGG CCGCGGCTCC CGGGTGTCCC AGGCCCCGCC GGTGCGCAGA GCATGGCGGG 60  
 TGCGGGCCCC AAGCGGCGCG CGCTAGCGGC GCCGGCGGCC GAGGAGAAGG AAGAGGCGCG 120  
 GGAGAAGATG CTGGCCGCCA AGAGCGCGGA CGGCTCGGCG CCGGCAGGCG AGGGCGAGGG 180  
 CGTGACCTCG CAGCGGAACA TCACGCTGCT CAACGCGGTG GCCATCATCG TGGGGACCAT 240  
 TATCGGCTCG GGCATCTTCG TGACGCCAC GGGCGTGCTC AAGGAGGCAG GCTCGCCGGG 300  
 GCTGGCGCTG GTGGTGTGGG CCGCGTGCGG CGTCTTCTCC ATCGTGGGCG CGCTCTGCTA 360  
 CGCGGAGCTC GGCACCACCA TCTCCAAATC GGGCGGCGAC TACGCTACA TGCTGGAGGT 420  
 CTACGGCTCG CTGCCCCCCT TCCTCAAGCT CTGGATCGAG CTGCTCATCA TCCGGCCTTC 480  
 ATCGCAGTAC ATCGTGCCCC TGGTCTTCGC CACCTACCTG CTCAAGCCGC TCTTCCCCAC 540  
 CTGCCCCGTG CCCGAGGAG CAGCCAAGCT CGTGGCCTGC CTCTGCGTGC TGCTGCTCAC 600  
 GGCCGTGAAC TGCTACAGCG TGAAGGCCGC CACCCGGGTC CAGGATGCCT TTGCGCGCGC 660  
 CAAGCTCCTG GCCCTGGCCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GAAAGGGTGA 720  
 TGTGTCCAAT CTAGATCCCA ACTTCTCATT TGAAGGACCA AACTGGATG TGGGGAACAT 780  
 TGTGCTGGCA TTATACAGCG GCCTCTTTGC CTATGGAGGA TGGAACTACT TGAATTTCTG 840  
 CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCTG GCCATCATCA TCTCCCTGCC 900  
 CATCGTGACG CTGGTGACG TGCTGACCAA CTTGGCTTAC TCCACACCC TGTCCACCGA 960  
 GCAGATGCTG TCGTCCGAGG CCGTGGCGGT GAGCTTCGGG AACTATCACC TGGGCGTCAT 1020  
 GTCCTGGATC ATCCCCGTCT TCGTGGGCCT GTCCTGCTTC GGCTCCGTCA ATGGGTCCCT 1080  
 GTTCACATCC TCCAGGCTCT TCTTCGTGGG GTCCCGGGAA GGCCACCTGC CCTCCATCCT 1140  
 CTCCATGATC CACCCACAGC TCTCACCCC CGTGCCGTCC CTCGTGTCA CGTGTGTGAT 1200  
 GACGCTGCTC TACGCTTCT CCAAGGACAT CTTCTCCGTC ATCAACTTCT TCAGCTTCTT 1260  
 CAACTGGCTC TGGCTGGCCC TGGCCATCAT TGGCTGCGCC ACAGAAAGCC 1320  
 TGAGCTTGAG CGGCCATCA AGGTGAACCT GGCCCTGCCT GTGTTCTTCA TCCTGGCCTG 1380  
 CCTCTTCTG ATCGCGTCT CTTCTGGAA GACACCCGTG GAGTGTGGCA TCGGCTTCA 1440  
 CATCATCCTC AGCGGGCTGC CCGTCTACTT CTTCCGGGTC TGGTGGAAAA ACAAGCCCAA 1500  
 GTGGCTCCTC CAGGGCATCT TCTCCACGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGGT 1560



CCCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

Seq ID NO: 543 Protein sequence  
Protein Accession #: XP\_035292.2

1	11	21	31	41	51	
MAGAGPKRRA	LAAPAAEEKE	EAREKMLAAK	SADGSAPAGE	GEGVTLQRNI	TLLNGVAIIV	60
GTIIGSGIFV	TPTGVLKEAG	SPGLALVWVA	ACGVFSIVGA	LCYAEELGTTI	SKSGGDYAYM	120
LEVYGSPLPAF	LKLWIELLII	RPSSQYIVAL	VFATYLLKPL	FPTCPVPEEA	AKLVACLCVL	180
LLTAVNCYSV	KAATRVQDAF	AAAKLLALAL	IILLGFVQIG	KGDVSNLDPN	FSFEGTKLDV	240
GNIVLALYSG	LFAYGGWNYL	NFVTEEMINP	YRNLPLAIII	SLPIVTLVYV	LTNLAYFTTL	300
STEQMLSSSEA	VAVDFGNYHL	GVMSWIIPVF	VGLSCFGSVN	GSFLTSSRLF	FVGSREGHLP	360
SILSMIHPQL	LTPVPSLVFT	CVMTLLYAFS	KDIFSVINFF	SFFNWLCVAL	AIIGMIWLRH	420
RKPELERPIK	VNLALPVFFI	LACLEFLIAVS	FWKTPVECGI	GFTIILSGLP	VYFFGVWWKN	480
KPKWLLQGIF	STTVLCQKLM	QVVPQET				

Seq ID NO: 544 DNA sequence  
Nucleic Acid Accession #: NM\_005268.1  
Coding sequence: 168..989

1	11	21	31	41	51	
TAAAAAGCAA	AAGAATTCGC	GGCCGCGTCG	ACACGGGCTT	CCCCGAAAAC	CTTCCCCGCT	60
TCTGGATATG	AAATTCAAGC	TGCTTGCTGA	GTCTATTGTC	CGGCTGCTGG	GAGCCAGGAG	120
AGCCCTGAGG	AGTAGTCACT	CAGTAGCAGC	TGACGCGTGG	GTCCACCATG	AACTGGAGTA	180
TCTTTGAGGG	ACTCCTGAGT	GGGTCAACA	AGTACTCCAC	AGCCTTTGGG	CGCATCTGGC	240
TGTCTCTGGT	CTTCATCTTC	CGCGTGCTGG	TGTACCTGGT	GACGGCCGAG	CGTGTGTGGA	300
GTGATGACCA	CAAGGACTTC	GACTGCAATA	CTCGCCAGCC	CGGCTGCTCC	AACGTCTGCT	360
TTGATGAGTT	CTTCCCTGTG	TCCCATGTGC	GCCTCTGGGC	CCTGCAGCTT	ATCCTGGTGA	420
CATGCCCTC	ACTGCTCGTG	GTCAATGACG	TGGCCTACCG	GGAGGTTTCA	GAGAAGAGGC	480
ACCGAGAAGC	CCATGGGGAG	AACAGTGGGC	GCCTCTACCT	GAACCCCGGC	AAGAAGCGGG	540
GTGGGCTCTG	GTGGACATAT	GTCTGCAGCC	TAGTGTTCAC	GGCGAGCGTG	GACATCGCCT	600
TTCTCTATGT	GTTCCACTCA	TTCTACCCCA	AATATATCCT	CCCTCCTGTG	GTCAAGTGCC	660
ACGCAGATCC	ATGTCCCAAT	ATAGTGGACT	GCTTCATCTC	CAAGCCCTCA	GAGAAGAACA	720
TTTTACCCCT	CTTCATGGTG	GCCACAGCTG	CCATCTGCAT	CCTGTCAAC	CTCGTGGAGC	780
TCATCTACCT	GGTGAGCAAG	AGATGCCACG	AGTGCCTGGC	AGCAAGGAAA	GCTCAAGCCA	840
TGTGCACAGG	TCATCACCCC	CACGGTACCA	CCTCTTCCTG	CAAACAAGAC	GACCTCCTTT	900
CGGGTGACCT	CATCTTTCTG	GGCTCAGACA	GTATCTCTCC	TCTCTTACCA	GACCGCCCCC	960
GAGACCATGT	GAGAAAACC	ATCTTGTGAG	GGGCTGCCTG	GACTGGTCTG	GCAGGTTGGG	1020
CCTGGATGGG	GAGGCTCTAG	CATCTCTCAT	AGGTGCAACC	TGAGAGTGGG	GGAGCTAAGC	1080
CATGAGGTAG	GGGCAGGCAA	GAGAGAGGAT	TCAGACGCTC	TGGGAGCCAG	TTCTAGTCC	1140
TCAACTCCAG	CCACCTGCCC	CAGCTCGACG	GCACTGGGCC	AGTTCCTCCCT	CTGCTCTGCA	1200
GCTCGGTTTC	CTTTTCTAGA	ATGGAATAG	TGAGGGCCAA	TGC		

Seq ID NO: 545 Protein sequence  
Protein Accession #: NP\_005259.1

1	11	21	31	41	51	
MNWSIFEGLL	SGVNKYSTAF	GRIWLSLVFI	FRVLVYLVT	ERVWSDDHKD	FDCNTRQPGC	60
SNVCFDEFFP	VSHVRLWALQ	LILVTCPSLH	VVMHVAYREV	QEKRHREAHG	ENSGRLYLNP	120
GKKRGGWLWT	YVCSLVFKAS	VDIAFLYVPH	SFYPKYILPP	VVKCHADPCP	NIVDCFISKP	180
SEKNIFTLFM	VATAAICILL	NLVELIYLV	KRCHECLAAR	KAQAMCTGHH	PHGTTSSCKQ	240
DDLKSGDLIF	LGSDSHPLLL	PDRPRDHVKK	TIL			

Seq ID NO: 546 DNA sequence  
Nucleic Acid Accession #: NM\_002391.1  
Coding sequence: 26..457

1	11	21	31	41	51	
CGGGCGAAGC	AGCGCGGGCA	GCGAGATGCA	GCACCGAGGC	TTCCTCCTCC	TCACCTCCT	60
CGCCCTGCTG	GCGCTCACCT	CCGCGGTGCG	CAAAAAGAAA	GATAAGGTGA	AGAAGGGCGG	120
CCCGGGGAGC	GAGTGCCTG	AGTGGGCTG	GGGGCCCTGC	ACCCCCAGCA	GCAAGGATTG	180
CGGCGTGGGT	TTCCGCGAGG	GCACCTGCGG	GGCCCAGACC	CAGCGCATCC	GGTGCAGGGT	240
GCCCTGCAAC	TGGAAGAAGG	AGTTTGGAGC	CGACTGCAAG	TACAAGTTTG	AGAACTGGGG	300
TGCGTGTGAT	GGGGGCACAG	GCACCAAAGT	CCGCCAAGGC	ACCCTGAAGA	AGGCGCGCTA	360
CAATGCTCAG	TGCCAGGAGA	CCATCCGCGT	CACCAAGCCC	TGCACCCCCA	AGACCAAAGC	420
AAAGGCCAAA	GCCAAGAAAG	GGAAGGGAAA	GGACTAGACG	CCAAGCCTGG	ATGCCAAGGA	480
GCCCCTGGTG	TCACATGGGG	CCTGGCCACG	CCCTCCCTCT	CCCAGGCCCG	AGATGTGACC	540
CACCAAGTGC	TTCTGTCTGC	TCGTAGTCTG	TAATCAATCA	TGCCCTGCCT	TGTCCCTCTC	600
ACTCCCAGC	CCCACCCCTA	AGTGCCCAAA	GTGGGGAGGG	ACAAGGGATT	CTGGGAAGCT	660
TGAGCCTCCC	CCAAAGCAAT	GTGAGTCCCA	GAGCCCGCTT	TTGTTCTTCC	CCACAATTCC	720
ATTACTAAGA	AACACATCAA	ATAAAGTAC	TTTTTCCCCC	CAATAAAAGC	TCTTCTTTTT	780
TAATAT						

Seq ID NO: 547 Protein sequence  
Protein Accession #: NP\_002382.1

1	11	21	31	41	51	
MQHRGFLLLT	LLALLALTS	VAKKKDKVKK	GGPGSECAEW	AWGPCTPSSK	DCGVGFREGT	60
CGAQTQIRIC	RVPCNWKKEF	GADCKYKFEN	WGACDGGTGT	KVRQGTLLKA	RYNAQCQETI	120
RVTKPCTPKT	KAKAKAKKKG	GKD				

Seq ID NO: 548 DNA sequence



Nucleic Acid Accession #: NM\_006783.1  
Coding sequence: 1..786

```
5 1 11 21 31 41 51
| | | | |
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAAACACTC CACCAGCATC 60
GGGAAGGTGT GGATCACAGT CATCTTTATT TTCCGAGTCA TGATCCTAGT GGTGGCTGCC 120
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACCGGGATGC 180
AAAAATGTGT GCTATGACCA CTTTTCCTCG GTGTCCCACA TCCGGCTGTG GGCCCTCCAG 240
10 CTGATCTTCG TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCCTA CTACAGGCAC 300
GAAACCACTC GCAAGTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360
ATTAAAAAGC ACAAGTTCG GATAGAGGGG TCGCTGTGGT GGACGTACAC CAGCAGCATC 420
TTTTTCCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACAATGGG 480
TACCACCTGC CCTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCAACCT TGTGACTGC 540
15 TTTATTTCTA TGCCCAACAGA GAAGACCGTG TTTACCATT TTTATGATTG TGCGTCTGTG 600
ATTTGCATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTTAGG 660
AGATCAAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720
CAGAATGAAA TGAATGAGCT GATTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780
20 AGCTAA
```

Seq ID NO: 549 Protein sequence  
Protein Accession #: NP\_006774.1

```
25 1 11 21 31 41 51
| | | | |
MDWGTLHTFI GGVNKHSTSI GKWITVIFI FRVMILVVAA QEVWGDEQED FVCNTLQPGC 60
KNVCYDHFPP VSHRLWALQ LIFVSTPALL VAMHVAYYRH ETTRKFRGE KRNDFKDIED 120
IKKHKVRIG SLWWTYTSSI FFRIFEEAAF MYVFYFLYNG YHLPWVLKCG IDPCPNLVDC 180
30 FISRPTEKTV FTIFMISASV ICMLLNVAEL CYLLLKVCFR RSKRAQTQKN HPNHALKESK 240
QNMENELISD SGQNAITGFP S
```

Seq ID NO: 550 DNA sequence  
Nucleic Acid Accession #: NM\_002571.1  
Coding sequence: 99..587

```
35 1 11 21 31 41 51
| | | | |
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCTCCTGC 60
TCACCTTGGG CGTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120
40 AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CTGGCACTC CATGGCCATG GCGACCAACA 180
ACATCTCCCT CATGGCGACA CTGAAGGCCC CTCTGAGGGT CCACATCACC TCACTGTTGC 240
CCACCCCGGA GGACAACCTG GAGATCGTTC TGCACAGATG GGAGAACAAC AGCTGTGTTG 300
AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG 360
45 TGGCGAACGA GGCCACGCTG CTCGATACTG ACTACGACAA TTTCTGTGTT CTCTGCCTAC 420
AGGACACCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCCTGGTGG 480
AGGACGATGA GATCATGCAG GGATTCATCA GGGCTTTCAG GCCCTGCCG AGGCACCTAT 540
GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCCGTGCCG TTTCTAGCTC ACCTCCGCCT 600
CCAGGAAGAC CAGACTCCCA CCCTTCCACA CCTCCAGAGC AGTGGGACTT CCTCCTGCCC 660
50 TTTCAAAGAA TAACCACAGC TCAGAAAGAG ATGACGTGGT CATCTGTGTC GCCATCCCCT 720
TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCCTGGGGGC AGAGTCTCTG 780
GCAGAGGTTA TTAATAAACC CTTGGAGCAT G
```

Seq ID NO: 551 Protein sequence  
Protein Accession #: NP\_002562.1

```
55 1 11 21 31 41 51
| | | | |
MDIPQTKQDL ELPKLAGTWH SMAMATNNIS LMATLKAPLR VHITSLLP TP EDNLEIVLHR 60
WENNSCVEKK VLGEKTGNPK KFKINYTVAN EATLLD TDYD NFLFLCLQDT TTP IQSMMCQ 120
60 YLARVLVEDD EIMQGFIRAF RPLPRHLWYL LDLKQMEEP RF
```

Seq ID NO: 552 DNA sequence  
Nucleic Acid Accession #: NM\_006500.1  
Coding sequence: 27..1967

```
65 1 11 21 31 41 51
| | | | |
ACTTGCGTCT CGCCCTCCGG CCAAGCATGG GGCTTCCCAG GCTGGTCTGC GCCTTCTTGC 60
TCGCCGCCTG CTGCTGCTGT CCTCGCGTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
70 CGCCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCTCTCTCC 180
AGTCCCAAGG CAACCTCAGC CATGTCGACT GGTTTTCTGT CCACAAGGAG AAGCGGACGC 240
TCATCTTCCG TGTGCGCCAG GGCCAGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGGC 300
TCAGCCTCCA GGACAGAGGG GCTACTCTGG CCCTGACTCA AGTCACCCCC CAAGACGAGC 360
GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCCGCG 420
75 TCTACAAAGC TCCGGAGGAG CCAAACATCC AGGTCAACCC CCTGGGCATC CCTGTGAACA 480
GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGGTACCCC ATTCCTCAAG 540
TCATCTGGTA CAAGAATGGC CGGCTCTGTA AGGAGGAGAA GAACCGGGTC CACATTCAGT 600
CGTCCCAGAC TGTGGAGTCG AGTGGTTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC 660
TGGTTAAAGA AGACAAAGAT GCCCAGTTT ACTGTGAGCT CAACTACCGG CTGCCAGTG 720
80 GGAACCACAT GAAGGAGTCC AGGGAAGTCA CCGTCCCTGT TTTCTACCG ACAGAAAAAG 780
TGTGGCTGGA AGTGGAGCCC GTGGGAATGC TGAAGGAAGG GGACCGCGTG GAAATCAGGT 840
GTTTGGCTGA TGGCAACCCT CCACCACACT TCAGCATCAG CAAGCAGAAC CCCAGCACCA 900
GGGAGGCAGA GGAAGAGACA ACCAAGCAGA CCGGGTCTCT GGTGCTGGAG CCTGCCCGGA 960
85 AGGAACACAG TGGGCGCTAT GAATGTCAGG CCTGGAACCT GGACACCATG ATATCGCTGC 1020
TGAGTGAACC ACAGGAAC TA CTGGTGAAC ATGTGTCTGA CGTCCGAGTG AGTCCCGCAG 1080
CCCCTGAGAG ACAGGAAGGC AGCAGCCTCA CCCTGACCTG TGAGGCAGAG AGTAGCCAGG 1140
ACCTCGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAGGT GCTGGAAAGG GGCCTGTGTC 1200
```



	TTCAGTTGCA	TGACCTGAAA	CGGGAGGCAG	GAGGCGGCTA	TCGCTGCGTG	GCGTCTGTGC	1260
	CCAGCATAAC	CGGCCTGAAC	CGCACACAGC	TGGTCAAGCT	GGCCATTTTT	GGCCCCCTT	1320
	GGATGGCATT	CAAGGAGAGG	AAGGTGTGGG	TGAAAGAGAA	TATGGTGTG	AATCTGTCTT	1380
5	GTGAAGCGTC	AGGGCACCCC	CGGCCCCACCA	TCTCCTGGAA	CGTCAACGGC	ACGGCAAGTG	1440
	AACAAGACCA	AGATCCACAG	CGAGTCCTGA	GCACCCTGAA	TGTCTCGTG	ACCCCGGAGC	1500
	TGTTGGAGAC	AGGTGTTGAA	TGCACGGCCT	CCAACGACCT	GGGCAAAAAC	ACCAGCATCC	1560
	TCTTCCTGGA	GCTGGTCAAT	TTAACCACCC	TCACACCAGA	CTCCAACACA	ACCACTGGCC	1620
	TCAGCACTTC	CACCTGCCAGT	CCTCATACCA	GAGCCAACAG	CACCTCCACA	GAGAGAAAGC	1680
10	TGCCGGAGCC	GGAGAGCCGG	GGCGTGGTCA	TCGTGGCTGT	GATTGTGTGC	ATCCTGGTCC	1740
	TGGCGGTGCT	GGGCGCTGTC	CTCTATTTCC	TCTATAAGAA	GGGCAAGCTG	CCGTGCAGGC	1800
	GCTCAGGGAA	GCAGGAGATC	ACGCTGCCCC	CGTCTCGTAA	GACCGAACTT	GATGTTGAAG	1860
	TTAAGTCAGA	TAAGCTCCCA	GAAGAGATGG	GCCTCCTGCA	GGGCAGCAGC	GGTGACAAGA	1920
	GGGCTCCGGG	AGACAGGGGA	GAGAAATACA	TCGATCTGAG	GCATTAGCCC	CGAATCACTT	1980
15	CAGCTCCCTT	CCCTGCCTGG	ACCATTCCCA	GCTCCCTGCT	CACCTCTCTC	TCAGCCAAAG	2040
	CCTCCAAAGG	GACTAGAGAG	AAGCCTCCTG	CTCCCCTCAC	CTGCACACCC	CCTTTTCAGAG	2100
	GGCCACTGGG	TTAGGACCTG	AGGACCTCAC	TTGGCCCTGC	AAGCCGCTTT	TCAGGGACCA	2160
	GTCCACCACC	ATCTCCTCCA	CGTTGAGTGA	AGCTCATCCC	AAGCAAGGAG	CCCCAGTCTC	2220
	CCGAGCGGGT	AGGAGAGTTT	CTTGCAAGAC	GTGTTTTTTC	TTTACACACA	TTATGGCTGT	2280
20	AAATACCTGG	CTTCCTGCCAG	CAGCTGAGCT	GGGTAGCCTC	TCTGAGCTGG	TTTCTGCCC	2340
	CAAAGGCTGG	CTTCCACCAT	CCAGGTGCAC	CACTGAAGTG	AGGACACACC	GGAGCCAGGC	2400
	GCCTGCTCAT	GTTGAAGTGC	GCTGTTTACA	CCCGCTCCGG	AGAGCACCCC	AGCGGCATCC	2460
	AGAAGCAGCT	GCAGTGTGTC	TGCCACCACC	CTCCTGCTCG	CCTCTTCAAA	GTCTCCTGTG	2520
	ACATTTTTTC	TTTGGTCAGA	AGCCAGGAAC	TGGTGTCAAT	CCTTAAAGA	TACGTGCCGG	2580
25	GGCCAGGTGT	GGTGGTCCAC	GCCTGTAATC	CCAGCACTTT	GGGAGGCCGA	GGCGGGCGGA	2640
	TCACAAAGTC	AGGACGAGAC	CATCCTGGCT	AACACGGTGA	AACCCTGTCT	CTACTAAAAA	2700
	TACAAAAAAA	AATTAGCTAG	GCGTAGTGCT	TGGCACCTAT	AGTCCCAGCT	ACTCGGAAGG	2760
	CTGAAGCAGG	AGAATGGTAT	GAATCCAGGA	GGTGGAGCTT	GCAGTGAGCC	GAGACCGTGC	2820
	CACTGCACTC	CAGCCTGGGC	AACACAGCGA	GACTCCGTCT	CGAGGAAAAA	AAAAGAAAAG	2880
30	ACGCGTACCT	GCGGTGAGGA	AGCTGGGCGC	TGTTTTCGAG	TTCAGGTGAA	TTAGCCTCAA	2940
	TCCCCGTGTT	CACCTGCTCC	CATAGCCCTC	TTGATGGATC	ACGTAAACT	GAAAGGCAGC	3000
	GGGGAGCAGA	CAAAGATGAG	GTCTACACTG	TCCTTCATGG	GGATTAAAGC	TATGGTTATA	3060
	TTAGCACCAA	ACTTCTACAA	ACCAAGCTCA	GGGCCCCAAC	CCTAGAAGGG	CCCAAATGAG	3120
	AGAATGGTAC	TTAGGGATGG	AAAACGGGGC	CTGGCTAGAG	CTTCGGGTGT	GTGTGTCTGT	3180
35	CTGTGTGTAT	GCATACATAT	GTGTGTATAT	ATGGTTTTGT	CAGGTGTGTA	AATTTGCAAA	3240
	TTGTTTCCTT	TATATATGTA	TGTATATATA	TATATGAAAA	TATATATATA	TATGAAAAAT	3300
	AAAGCTTAAT	TGTCCCAGAA	AATCATACAT	TGCTTTTTTA	TTCTACATGG	GTACCACAGG	3360
	AACCTGGGGG	CCTGTGAAAC	TACAACCAAA	AGGCACACAA	AACCGTTTCC	AGTTGGCAGC	3420
	AGAGATCAGG	GGTTACCTCT	GCTTCTGAGC	AAATGGCTCA	AGCTCTACCA	GAGCAGACAG	3480
40	CTACCCTACT	TTTCAGCAGC	AAAACCTCCC	GTATGACGCA	GCACGAAGGG	CCTGGCAGGC	3540
	TGTTAGCAGG	AGCTATGTCC	CTTCCTATCG	TTTCCGTCCA	CTT		

Seq ID NO: 553 Protein sequence  
Protein Accession #: NP\_006491.1

45	1	11	21	31	41	51	
	GLPRLVCAFL	LAACCCCPRV	AGVPGEAEQF	APELVEVEVG	STALLKCGLS	QSQGNLSHVD	60
	WFSVHKEKRT	LIFRVRQGGQ	QSEPGYEYQR	LSLQDRGATL	ALTQVTPQDE	RIFLCQGKRP	120
50	RSQEYRIQLR	VYKAPEEPNI	QVNPLGIPVN	SKEPEEVATC	VGRNGYPIPIQ	VIWYKNGRPL	180
	KEEKNRVHIQ	SSQTVESSGL	YTLQSLKAQ	LVKEDKDAQF	YCELNYRLPS	GNHMKESREV	240
	TVPVFYPTFK	VWLEVEPVGM	LKEGDRVEIR	CLADGNPPPH	FSISKQNPST	REAEETND	300
	NGVLVLEPAR	KEHSGRYECQ	AWNLDTMISL	LSEPEQLLVN	YVSDVRVSPA	APERQEGSSL	360
	TLTCEAESSQ	DLEFQWLREE	TDQVLERGPV	LQLHDLKREA	GGGYRCVASV	PSIPGLNRTQ	420
55	LVKLAIFFGP	WMAFKERKVV	VKENMVLNLS	CEASGHPRPT	ISWNVNGTAS	EQDQDPQRVL	480
	STLNLVLTPE	LLETGVECTA	SNDLGKNTSI	LFLELVNLTT	LTPDSNNTTG	LSTSTASPHT	540
	RANSTSTERK	LPEPESRGVV	IVAVIVCILV	LAVLGAVLYF	LYKKGKLPCR	RSQKQBITLP	600
	PSRKTELVEE	VKSDKLPEEM	GLLQSSGDK	RAPGDQGEKY	IDLRH		

Seq ID NO: 554 DNA sequence  
Nucleic Acid Accession #: NM\_003183.3  
Coding sequence: 165..2639

65	1	11	21	31	41	51	
	TCGAGCCTGG	CGGTAGAATC	TTCCCAGTAG	GCGGCGCGGG	AGGGAAAAGA	GGATTGAGGG	60
	GCTAGGCCGG	GCGGATCCCG	TCCTCCCCCG	ATGTGAGCAG	TTTTCCGAAA	CCCCGTCAGG	120
	CGAAGGCTGC	CCAGAGAGGT	GGAGTCCGTA	GCGGGGCGCG	GAACATGAGG	CAGTCTCTCC	180
	TATTCCTGAC	CAGCGTGGTT	CCTTTCGTGC	TGGCGCGCGG	ACCTCCGGAT	GACCCGGGCT	240
70	TCGGCCCCCA	CCAGAGACTC	GAGAAGCTTG	ATTCTTTGCT	CTCAGACTAC	GATATTCTCT	300
	CTTTATCTAA	TATCCAGCAG	CATTCGTAA	GAAAAAGAGA	TCTACAGACT	TCAACACATG	360
	TAGAAACACT	ACTAATTTT	TCAGCTTTGA	AAAGGCATTT	TAAATTATAC	CTGACATCAA	420
	GTAAGTGAAC	TTTTTCACAA	AATTTCAAGG	TCGTGGTGGT	GGATGGTAAA	AACGAAAGCG	480
	AGTACACTGC	AAAATGGCAG	GACTTCTTCA	CTGGACACGT	GGTTGGTGAG	CCTGACTCTA	540
75	GGGTTCTAGC	CCACATAAGA	GATGATGATG	TTATAATCAG	AATCAACACA	GATGGGGCCG	600
	AATATAACAT	AGAGCCACTT	TGGAGATTGT	TTAAGTATAC	CAAAGACAAA	AGAATGTTAG	660
	TTTATAAATC	TAGAATATC	AAGAAATTTT	CACGTTTGCA	GTCTCCAAAA	GTGTGTGGTT	720
	ATTTAAAGT	GGATAATGAA	GAGTTGTCTC	CAAAAGGGTT	AGTAGACAGA	GAACCACTG	780
	AAGAGCTTGT	TCATCGAGTG	AAAAGAAGAG	CTGACCCAGA	TCCCATGAAG	AACACGTGTA	840
80	AATTATTGGT	GGTAGCAGAT	CATCGCTTCT	ACAGATACAT	GGGCAGAGGG	GAAGAGAGTA	900
	CAACTACAAA	TTACTTAATA	GAGCTAATTG	ACAGAGTTGA	TGACATCTAT	CGGAACACTT	960
	CATGGGATAA	TGCAGGTTTT	AAAGGCTATG	GAATACAGAT	AGAGCAGATT	CGCATCTCTA	1020
	AGTCTCCACA	AGAGGTAAAA	CCTGGTGAAA	AGCACTACAA	CATGGCAAAA	AGTTACCCAA	1080
	ATGAAGAAAA	GGATGCTTGG	GATGTGAAGA	TGTTGCTAGA	GCAATTTAGC	TTTGATATAG	1140
85	CTGAGGAAGC	ATCTAAAGTT	TGCTTGGCAC	ACCTTTTCAC	ATACCAAGAT	TTTGATATGG	1200
	GAACCTTTGG	ATTAGCTTAT	GTTGGCTCTC	CCAGAGCAAA	CAGCCATGGA	GGTTTGTGTC	1260
	CAAAGGCTTA	TTATAGCCCA	GTTGGGAAGA	AAAATATCTA	TTTGAATAGT	GGTTTGACGA	1320
	GCACAAAGAA	TTATGGTAAA	ACCATCCTTA	CAAAGGAAGC	TGACCTGGTT	ACAACCTCATG	1380



	AATTGGGACA	TAATTTTGGG	GCAGAACATG	ATCCGGATGG	TCTAGCAGAA	TGTGCCCCGA	1440
	ATGAGGACCA	GGGAGGGAAA	TATGTCATGT	ATCCCATAGC	TGTGAGTGGC	GATCACGAGA	1500
	ACAATAAGAT	GTTTTCAAAC	TGCAGTAAAC	AATCAATCTA	TAAGACCATT	GAAAGTAAGG	1560
5	CCCAGGAGTG	TTTTCAAGAA	CGCAGCAATA	AAGTTTGTGG	GAACCTGAGG	GTGGATGAAG	1620
	GAGAAGAGTG	TGATCCTGGC	ATCATGTATC	TGAACAACGA	CACCTGCTGC	AACAGCGACT	1680
	GCACGTTGAA	GGAAGGTGTC	CAGTGCAGTG	ACAGGAACAG	TCCTTGCTGT	AAAAACTGTC	1740
	AGTTTGTAGAC	TGCCCAGAAG	AAGTGCCAGG	AGGCGATTAA	TGCTACTTGC	AAAGGCGTGT	1800
	CCTACTGCAC	AGGTAATAGC	AGTGAGTGCC	CGCCTCCAGG	AAATGCTGAA	AATGACACTG	1860
10	TTTGCTTGGG	TCTTGGCAAG	TGTAAGGATG	GGAAATGCAT	CCCTTCTGTC	GAGAGGGAAC	1920
	AGCAGCTGGA	GTCCTGTGCA	TGTAATGAAA	CTGACAACCTC	CTGCAAGGTG	TGCTGCAGGG	1980
	ACCTTTCTGG	CCGCTGTGTG	CCCTATGTCT	ATGCTGAACA	AAAGAACTTA	TTTTTGAGGA	2040
	AAGGAAAGCC	CTGTACAGTA	GGATTTTGTG	ACATGAATGG	CAAATGTGAG	AAACGAGTAC	2100
	AGGATGTAAT	TGAACGATTT	TGGGATTTCA	TTGACCAGCT	GAGCATCAAT	ACTTTTGGAA	2160
15	AGTTTTTAGC	AGACAACATC	GTTGGGTCTG	TCCTGGTTTT	CTCCTTGATA	TTTTGGATTG	2220
	CTTTCAGCAT	TCTTGTCCAT	TGTGTGGATA	AGAAATTGGA	TAAACAGTAT	GAATCTCTGT	2280
	CTCTGTTTCA	CCCCAGTAAC	GTCGAAATGC	TGAGCAGCAT	GGATTCTGCA	TCGGTTCGCA	2340
	TTATCAAACC	CTTTCCTGCG	CCCCAGACTC	CAGGCCGCTT	GCAGCCTGCC	CCTGTGATCC	2400
	CTTCGGCGCC	AGCAGCTCCA	AAACTGGACC	ACCAGAGAAT	GGACACCATC	CAGGAAGACC	2460
20	CCAGCACAGA	CTCCCATATG	GACGAGGATG	GGTTTGAGAA	GGACCCCTTC	CCAAATAGCA	2520
	GCACAGCTGC	CAAGTCATTT	GAGGATCTCA	CGGACCATCC	GGTCGCCAGA	AGTGAAAAGG	2580
	CTGCCTCCTT	TAAACTGCAG	CGTCAGATCA	TGTTTAAACAG	CAAAGAAAACA	GAGTGCTAAT	2640
	TTAGTTCTCA	GCTCTTCTGA	CTTAAGTGTG	CAAAATATTT	TTATAGATTT	GACCTACAAA	2700
	TCAATCACAG	CTTGTATTTT	GTGAAGACTG	GGAAGTGAAT	TAGCAGATGC	TGGTCATGTG	2760
25	TTTGAACCTC	CTGCAGGTAA	ACAGTCTCTT	TGTGGTTTGG	CCCTTCTCCT	TTTGAAAAGG	2820
	TAAGGTGAAA	GTGAATCTAC	TTATTTTGAG	GCTTTCAGGT	TTAGTTTTTT	AAAATATCTT	2880
	TTGACCTGTG	GTGCAAAAGC	AGAAAATACA	GCTGGATTGG	GTTATGAATA	TTTACGTTTT	2940
	TGTAAATTAA	TCTTTTATAT	TGATAACAGC	ACTGACTAGG	GAAATGATCA	GTTTTTTTTT	3000
	ATACACTGTA	ATGAACCGCT	GAATATGAAG	CATTTGGCAT	TTATTTGTGA	GAAAAGTGGA	3060
30	ATAGTTTTTT	TTTTTTTTTT	TTTTTTTTTG	CTTCAACTAA	AAACAAAGGA	GATAAATTTA	3120
	GTATACATTG	TATCTAAATT	GTGGGTCTAT	TTCTAGTTAT	TACCCAGAGT	TTTTATGTAG	3180
	CAGGGAAAAT	ATATATCTAA	ATTTAGAAAT	CATTTGGGTT	AATATGCTCT	TTCATAATTC	3240
	TAAGACTAAT	GCTCAGAACC	TAACCACTAC	CTTACAGTGA	GGGCTATACA	TGGTAGCCAG	3300
	TTGAATTTAT	GGAATCTACC	AACCTGTTTA	GGCCCTGATT	TGCTGGGCAG	TTTTTCTGTA	3360
35	TTTTATAAGT	ATCTTCATGT	ATCCTGTGTA	CTGATAGGGA	TACATGTCTT	AGAAAATTCA	3420
	CTATTGGCTG	GGAGTGGTGG	CTCATGCCTG	TAATCCCAGC	ACTTGGAGAG	GCTGAGGTTG	3480
	CGCCACTACA	CTCCAGCCTG	GGTGACAGAG	TGAGATCTGC	CTC		

Seq ID NO: 555 Protein sequence  
Protein Accession #: NP\_003174.2

40	1	11	21	31	41	51	
	MRQSLLFLTS	VVPFVLAPRP	PDDPGFGPHQ	RLEKLDLSLLS	DYDILSLSNL	QQHSVRKRDL	60
45	QTSTHVETLL	TFSALKRHFH	LYLTSSTERF	SQNFKVVVVD	GKNESEYTA	WQDFFTGHV	120
	GEPDSRVLAH	IRDDDVIIIR	NTDGAEYNH	PLWRFVNDTK	DKRMLVYKSE	DIKNVSRLOS	180
	PKVCGYLKVD	NEELLPKGLV	DREPPPELVH	RVKRRADPD	MKNCTKLLVV	ADHRFYRYMG	240
	RGEESTTINY	LIELIDRVDD	IYRNTSWDNA	GFKGYGIQIE	QIRILKSPQE	VKPGKHYNM	300
	AKSYPNEEKD	AWDVKMLLEQ	FSFDIAEEAS	KVCLAHLFY	QDFDMGTLGL	AYVGSPRANS	360
50	HGGVCPKAYY	SPVGKKNIYL	NSGLTSTKNY	GKTILTKEAD	LVTTHLGHN	FGAEHDPDGL	420
	AECAPNEDQG	GKYVMYPIAV	SGDHENNKMF	SNCSKQSIYK	TIESKAQECF	QERSNKVCGN	480
	SRVDEGEED	PGIMYLNNDT	CCNSDCTLKE	GVQCSDRNSP	CKKNCQFETA	QKKCQEAINA	540
	TKCGVSYCTG	NSSECPPPGN	AENDTVCLDL	GKCKDGKCI	FCEREQQLES	CACNETDNCS	600
	KVCCRDLSGR	CVPYVDAEQK	NLFRLKGPCK	TVGFCDMNGK	CEKRVQDVIE	RFWDFIDQLS	660
55	INTFGKFLAD	NIVGSVLVFS	LIFWIPFSIL	VHCVDKKLDK	QYESLSLFHP	SNVEMLSMMD	720
	SASVRIIKPF	PAPQTPGRLQ	PAPVIPSAPA	APKLDHQKMD	TIQEDPSTDS	HMDDEDGFEKD	780
	PPFNSSTA	AK SFEDLTDHPV	ARSEKAASFK	LQRQNRVNSK	ETEC		

Seq ID NO: 556 DNA sequence  
Nucleic Acid Accession #: NM\_021832.1  
Coding sequence: 164..2248

60	1	11	21	31	41	51	
	TCGAGCCTGG	CGGTAGAATC	TTCCCAGTAG	GCGGCGCGGG	AGGAAAAGAG	GATTGAGGGG	60
65	CTAGGCCCGG	CGGATCCCGT	CCTCCCCCGA	TGTGAGCAGT	TTTCCGAAAC	CCCGTCAGGC	120
	GAAGGCTGCC	CAGAGAGGTG	GAGTCGGTAG	CGGGGCCGGG	AACATGAGGC	AGTCTCTCCT	180
	ATTCTTGACC	AGCGTGGTTC	CTTTCGTGCT	GGCGCCGCGA	CCTCCGGATG	ACCCGGGCTT	240
	CGGCCCCCAC	CAGAGACTCG	AGAAGCTTGA	TTCTTTGCTC	TCAGACTACG	ATATTCTCTC	300
70	TTTATCTAAT	ATCCAGCAGC	ATTCCGTAAG	AAAAAGAGAT	CTACAGACTT	CAACACATGT	360
	AGAAACACTA	CTAACTTTTT	CAGCTTTGAA	AAGGCATTTT	AAATTATACC	TGACATCAAG	420
	TACTGAACGT	TTTTACAAA	ATTTCAAGGT	CGTGGTGGTG	GATGGTAAAA	ACGAAAGCGA	480
	GTACACTGTA	AAATGGCAGG	ACTTCTTCAC	TGGACACGTG	GTTGGTGAGC	CTGACTCTAG	540
	GGTTCTAGCC	CACATAAGAG	ATGATGATGT	TATAATCAGA	ATCAACACAG	ATGGGGCCGA	600
75	ATATAACATA	GAGCCACTTT	GGAGATTTGT	TAATGATACC	AAAGACAAA	GAATGTTAGT	660
	TTATAAATCT	GAAGATATCA	AGAATGTTTC	ACGTTTGCAG	TCTCCAAAAG	TGTGTGGTTA	720
	TTTAAAGTG	GATAATGAAG	AGTTGCTCCC	AAAAGGGTTA	GTAGACAGAG	AACCACCTGA	780
	AGAGCTTGTT	CATCGAGTGA	AAAGAAGAGT	TGACCCAGAT	CCCATGAAGA	ACAGTGTA	840
	ATTATTGGTG	GTAGCAGATC	ATCGCTTCTA	CAGATACATG	GGCAGAGGGG	AAGAGAGTAC	900
80	AACTACAAAT	TACTTAATAG	AGCTAATTGA	CAGAGTTGAT	GACATCTATC	GGAACACTTC	960
	ATGGGATAAT	GCAGGTTTTA	AAGGCTATGG	AATACAGATA	GAGCAGATTC	GCATTCTCAA	1020
	GTCTCCACAA	GAGGTAAAAC	CTGGTGAAAA	GCACTACAAC	ATGGCAAAAA	GTTACCCAAA	1080
	TGAAGAAAAG	GATGCTTGGG	ATGTGAAGAT	GTTGCTAGAG	CAATTTAGCT	TTGATATAGC	1140
	TGAGGAAGCA	TCTAAAGTTT	GCTTGGCACA	CCTTTTCA	TACCAAGATT	TTGATATGGG	1200
85	AACTCTTGGA	TTAGCTTATG	TTGGCTTCTC	CAGAGCAAA	AGCCATGGAG	GTGTTTGTCC	1260
	AAAGGCTTAT	TATAGCCAG	TTGGGAAGAA	AAATATCTAT	TTGAATAGTG	GTTTGACGAG	1320
	CACAAAGAAT	TATGGTAAAA	CCATCCTTAC	AAAGGAAGCT	GACCTGGTTA	CAACTCATGA	1380
	ATTGGGACAT	AATTTTGGAG	CAGAACATGA	TCCGGATGGT	CTAGCAGAA	GTGCCCCGAA	1440



5  
10  
15  
20  
25  
30  
35

TGAGGACCAG GGAGGGAAT ATGTCATGTA TCCCATAGCT GTGAGTGGCG ATCACGAGAA 1500  
CAATAAGATG TTTTCAAAC GCAGTAAACA ATCAATCTAT AAGACCATTG AAAGTAAGGC 1560  
CCAGGAGTGT TTTCAAGAAC GCAGCAATAA AGTTTGTGGG AACTCGAGGG TGGATGAAGG 1620  
AGAAGAGTGT GATCCTGGCA TCATGTATCT GAACAACGAC ACCTGCTGCA ACAGCGACTG 1680  
CACGTTGAAG GAAGGTGTCC AGTGCAGTGA CAGGAACAGT CCTTGCTGTA AAAACTGTCA 1740  
GTTTGAGACT GCCCAGAAGA AGTGCCAGGA GGCGATTAAT GCTACTTGCA AAGGCGTGTC 1800  
CTACTGCACA GGTAAATAGCA GTGAGTGGCC GCCTCCAGGA AATGCTGAAG ATGACACTGT 1860  
TTGCTTGGAT CTGGCAAGT GTAAGGATGG GAAATGCATC CCTTCTGCG AGAGGGAACA 1920  
GCAGCTGGAG TCCTGTGCAT GTAATGAAAC TGACAACTCC TGCAAGGTGT GCTGCAGGGA 1980  
CCTTCCGGC CGCTGTGTGC CCTATGTGCA TGCTGAACAA AAGAACTTAT TTTTGAGGAA 2040  
AGGAAAGCCC TGTACAGTAG GATTTTGTGA CATGAATGGC AAATGTGAGA AACGAGTACA 2100  
GGATGTAATT GAACGATTTT GGGATTTTCT TGACCAGCTG AGCATCAATA CTTTGGGAAA 2160  
GTTTTTAGCA GACAACATCG TTGGGTCTGT CCTGGTTTTT TCCTTGATAT TTTGGATTCC 2220  
TTTCAGCATT CTGTGCCATT GTGTGTAACG TCGAAATGCT GAGCAGCATG GATTCTGCAT 2280  
CGTTTCGCAT TATCAAACCC TTCTCTGCGC CCCAGACTCC AGGCCGCTG CAGCCTGCCC 2340  
CTGTGATCCC TTCGGCGCCA GCAGCTCCAA AACTGGACCA CCAGAGAATG GACACCATCC 2400  
AGGAAGACCC CAGCACAGAC TCACATATGG ACGAGGATGG GTTTGAGAAG GACCCCTTCC 2460  
CAAATAGCAG CACAGCTGCC AAGTCATTTG AGGATCTCAC GGACCATCCG GTCACCAGAA 2520  
GTGAAAAGGC TGCCTCCTTT AAAGTGCAGC GTCAGAATCG TGTGACAGC AAAGAAACAG 2580  
AGTGCTAATT TAGTTCTCAG CTCTTCTGAC TTAAGTGTGC AAAATATTTT TATAGATTTG 2640  
ACCTACAATC AATCACAGCT TATATTTTGT GAAGACTGGG AAGTGACTTA GCAGATGCTG 2700  
GTCATGTGTT TGAACCTTCT GCAGGTAAAC AGTTCTTGTG TGGTTTGGCC CTCTCCTTTT 2760  
TGAAAAGGTA AGGTGAAGGT GAATCTAGCT TATTTTGAGG CTTTCAGGTT TTAGTTTTTA 2820  
AAATATCTTT TGACCTGTGG TGCAAAGCA GAAAATACAG CTGGATTGGG TTATGAGTAT 2880  
TTACGTTTTT GTAAATTAAT CTTTATATAT GATAACAGGC ACTGACTAGG GAAATGATCA 2940  
GTTTTTTTTT ATACACTGTA ATGAACCGCT GAATATGAAG CATTTGCCAT TTATTTGTGA 3000  
GAAAAGTGGG ATAGTTTTTT TTTTTTTTTT TTTTTTTTGC CTTCAACTAA AAACAAAGGA 3060  
GATAAATTTA GTATACATTG TATCTAAAT GTGGGTCTAT TTCTAGTTAT TACCCAGAGT 3120  
TTTTATGTAG CAGGGAAAAT ATATATCTAA ATTTAGAAAT CATTGGGTT AATATGGCTC 3180  
TTCATAATTC TAAGACTAAT GCTCAGAAC TAACCACTAC CTTACAGTGA GGGCTATACA 3240  
TGGTAGCCAG TTGAATTTAT GGAATCTACC AACTGTTTAG GGCCCTGATT TGCTGGGCAG 3300  
TTTTTCTGTA TTTTATAAGT ATCTTCATGT ATCCCTGTTA CTGATAGGGA TACATGTCTT 3360  
AGAAAATTCA CTATTGGCTG GGAGTGGTGG CTCATGCCTG TAATCCAGC ACTTGGAGAG 3420  
3421 GCTGAGGTTG CGCCACTACA CTCCAGCCTG GGTGACAGAG TGAGATCTGC CTC

Seq ID NO: 557 Protein sequence  
Protein Accession #: NP\_068604.1

40  
45  
50

1 11 21 31 41 51  
| | | | |  
MRQSLFLTS VVPFVLAPRP PDDPGFGPHQ RLEKLDLSLS DYDILSLSN QHRSVRKRD 60  
QTSTHVETLL TFSALKRHF LYLTSSTERF SQNFVWVVD GKNESEYTVK WQDFFTGHV 120  
GEPDSRVLAH IRDDVIRI NTDGAEYNIE PLWRVNDTK DKRMLVYKSE DIKNVSRLOS 180  
PKVCGYLVVD NEELLPKGLV DREPPPELVH RVKRRADPDP MKNTCKLLV ADHRFYRYMG 240  
RGEESTTTNY LIELIDRVDD IYRNTSWDNA GFKGYGQIE QIRILKSPQE VKPGEKHYNM 300  
AKSYPNEEKD AWDVKMLLEQ FSPDIABEAS KVLALHFTY QDFDMGTLGL AYVGSPRANS 360  
HGGVCPKAYY SPVGKKNYIL NSGLTSTKNY GKTILTKEAD LVTTHLGHN FGAEHDPDGL 420  
AECAPNEDQG GKYVMPYIAV SGDHEHNKMF SNCSKQSIYK TIESKAQECF QERSNKVCGN 480  
SRVDEGEEDC PGIMYLNNDT CCNSDCTLKE GVQCSDRNSP CCKNCQFETA QKKCQEAINA 540  
TCKGVSYCTG NSSECPPPGN AEDDTVCLDL GKCKDGKCI FEREQQLES CACNETDNNSC 600  
KVCCRDLSGR CVPYVDAEQK NLFRLRGKPC TVGFCDMNGK CEKRVQDVIE RFWDFIDQLS 660  
INTFGKFLAD NIVGSVLVFS LIFWIPFSIL VHCV

Seq ID NO: 558 DNA sequence  
Nucleic Acid Accession #: NM\_004994.1  
Coding sequence: 20..2143

60  
65  
70  
75  
80  
85

1 11 21 31 41 51  
| | | | |  
AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCCTG GTCCTGGTGC TCCTGGTGCT 60  
GGGCTGCTGC TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTTGTGCTCT TCCCTGGAGA 120  
CCTGAGAACC AATCTCACCG ACAGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGTTA 180  
CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTTCT 240  
CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300  
GCGAACCCCA CGGTGCGGGG TCCAGAGATT CAAACCTTGT AGGGCGACCT 360  
CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTGCCGCG 420  
GGCGGTGATT GACGACGCCT TTGCCGCGC CTTGCACTG TGGAGCGCGG TGACGCGGCT 480  
CACCTTCACT CGCGTGTACA GCCGGGACGC AGACATCGTC ATCCAGTTTG GTGTGCGGGA 540  
GCACGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACG CCTTTCCTCC 600  
TGGCCCCGGC ATTACGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCTGGGCAA 660  
GGGCGTCTG GTTCCAATC GGTTTGAAA CGCAGATGGC GCGGCTGCC ACTTCCCTT 720  
CATCTTCGAG GGCCGCTCCT ACTCTGCTG CACCACCGAC GGTGCTCCG ACGGCTTGCC 780  
CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGCTTCT GCCCCAGCGA 840  
GAGACTCTAC ACCCGGGACG GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT 900  
CCAAGGCCAA TCCTACTCCG CCGCACCAC GGACGGTCGC TCCGACGGCT ACCGCTGGTG 960  
CGCCACCACC GCCAACTACG ACCGGGACAA GCTCTTCGGC TTCTGCCCGA CCCGAGCTGA 1020  
CTCGACGGTG ATGGGGGGCA ACTCGGCGGG GGAGCTGTGC GTCTTCCCT TCACTTTCCT 1080  
GGGTAAGGAG TACTCGACCT GTACCGCGA GATGGGCGCC TCTGGTGC 1140  
TACCACCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200  
TTTGTTCCTC GTGGCGGCGC ATGAGTTCGG CCACGCGCTG GGCTTAGATC ATTCCTCAGT 1260  
GCCGGAGGCG CTCATGTACC CTATGTACCG CTTCACTGAG GGGCCCCCT TGCATAAGGA 1320  
CGACGTGAAT GGCATCCGGC ACCTCTATGG TCCTCGCCCT GAACCTGAGC CACGGCCTCC 1380  
AACCACCACC ACACCGCAGC CCACGGCTCC CCCGACGGT TGGCCACCG GACCCCCAC 1440  
TGTCACCCCC TCAGAGCGCC CCACAGCTGG CCCCACAGT CCCCCCTCAG CTGGCCCCAC 1500  
AGGTCCCCCC ACTGCTGGCC CTTCTACGGC CACTACTGTG CTTTGTAGTC CGGTGGACGA 1560  
TGCTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAAACAGC TGTATTTGTT 1620  
CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGGG AGCCGGCCGC AGGCCCTT 1680



5  
10  
15

```

CCTTATCGCC GACAAGTGGC CCGCGCTGCC CCGCAAGCTG GACTCGGTCT TTGAGGAGCC 1740
GCTCTCCAAG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC 1800
GGTGCTGGGC CCGAGGCGTC TGGACAAGCT GGGCCTGGGA GCCGACGTGG CCCAGGTGAC 1860
CGGGGCCCTC CCGAGTGGCA GGGGGAAGAT GCTGCTGTTC AGCGGGCGGC GCCTCTGGAG 1920
GTTTCGACGTG AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT 1980
CCCCGGGGTG CCTTTGGACA CGCACGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG 2040
CCAGGACCGC TTCTACTGGC GCGTGAGTTC CCGGAGTGAG TTGAACCAGG TGGACCAAGT 2100
GGGCTACGTG ACCTATGACA TCCTGCAGTG CCCTGAGGAC TAGGGCTCCC GTCTTGCTTT 2160
GCAGTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCCGGATA 2220
CAAACTGGTA TTCTGTCTTG GAGGAAAGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC 2280
TCACCTTTGT TTTTGTGG AGTGTTCCTA ATAACTTGG ATTCTCTAAC CTTT

```

Seq ID NO: 559 Protein sequence  
Protein Accession #: NP\_004985.1

20  
25  
30

```

1 11 21 31 41 51
| | | | |
MSLWQPLVLV LLVLGCCFAA PRQRQSTLVL FPGDLRTNLT DRQLAEELYLY RYGYTRVAEM 60
RGESKSLGPA LLLLQKQLSL PETGELDSAT LKAMRTPRCG VPDLGRFQTF EGDLEKWHHHN 120
ITYWIQNYSE DLPRVIDDA FARAFALWSA VTPLTFTRVY SRDADIVIQF GVAEHGDSYF 180
FDGKDGLLAH AFPPGPGIQQ DAHFDDELW SLGKGVVVP RFGNADGAAC HFPFIEGERS 240
YSACTTDGRS DGLPWCSTTA NYDTDDRFGF CPSERLYTRD GNADGKPCQF PFIFQGSYS 300
ACTTDGRSDG YRWCATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTFLGKEYST 360
CTSEGRGDGR LWCATTSNFD SDKKWGFCDP QGYSLFLVAA HEFGHALGLD HSSVPEALMY 420
PMYRFTEGPP LHKDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAPPTVCPT GPPTVHPSER 480
PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFDKDGKYW 540
RFSEGRGSRP QGPFLIADKW PALPRKLDV FEPLSKKLF FFSGRQVWVY TGLSVLGP RR 600
LDKLGGLADV AQVTGALRSG RGKMLLFSGR RLWREDVKAQ MVDPRSASEV DRMFPGVPLD 660
THDVFQYREK AYFCQDRFYW RVSSRSELNQ VDQVGYVITYD ILQCPED

```

Seq ID NO: 560 DNA sequence  
Nucleic Acid Accession #: NM\_000213.1  
Coding sequence: 127..5385

35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

```

1 11 21 31 41 51
| | | | |
CGCCCCGCGC CTGCAGCCCC ATCTCCTAGC GGCAGCCAG GCGCGGAGGG AGCGAGTCCG 60
CCCCGAGGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GGCTGGCCGG GAGAGGGAGG 120
AAGAGGATGG CAGGGCCACG CCCAGCCCA TGGGCCAGGC TGCTCCTGGC AGCCTTGATC 180
AGCGTCAGCC TCTCTGGGAC CTTGGCAAAG AGGCCCCAGT GAAGAGTGC 240
ACGGAGTGTG TCCGTGTGGA TAAGGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC 300
CGGCGCTGCA ACACCCAGGC GGAGCTGCTG GCCGCGGGCT GCCAGCGGGA GAGCATCGTG 360
GTCATGGAGA GCAGCTTCCA AATCACAGAG GAGACCCAGA TTGACACCAC CCTGCGGCGC 420
AGCCAGATGT CCCCCCAAGG CCTGCGGGTC CGTCTGCGGC CCGGTGAGGA GCGGCATTTT 480
GAGCTGGAGG TGTTTGAGCC ACTGGAGAGC CCCGTGGACC TGTACATCCT CATGGACTTC 540
TCCAACCTCA TGTCCGATGA TCTGGCAAG CCAAGAAAGA TGGGGCAGAA CCTGGCTCGG 600
GTCCTGAGCC AGCTCACCAG CGACTACACT ATTGGATTTG GCAAGTTTGT GGACAAAGTC 660
AGCGTCCCGC AGACGGACAT GAGGCCTGAG AAGCTGAAGG AGCCCTGGCC CAACAGTGAC 720
CCCCCCTTCT CCTTCAAGAA CGTCATCAGC CTGACAGAAG ATGTGGATGA GTTCCGGAAT 780
AAACTGCAGG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CTGAGGGCGG CTTGCGATGCC 840
ATCCTGCAGA CAGCTGTGTG CACGAGGGAC ATTGGCTGGC GCCCGGACAG CACCCACCTG 900
CTGGTCTTCT CCACCGAGTC AGCCTTCCAT TATGGCGCAA CGTGCTGGCT 960
GGCATCATGA GCCGCAACGA TGAACCTGTC CACCTGGACA CCACGGGCAC CTACACCCAG 1020
TACAGGACAC AGGACTACCC GTCGGTGCCC ACCCTGGTGC GCCTGCTCGC CAAGCACAAC 1080
ATCATCCCCA TCTTTGCTGT CACCAACTAC TCCTATAGCT ACTACGAGAA GCTTCACACC 1140
TATTTCCCTG TCTCCTCACT GGGGGTGCTG CAGGAGGACT CGTCCAACAT CGTGGAGCTG 1200
CTGGAGGAGG CCTTCAATCG GATCCGCTCC AACCTGGACA TCCGGGCCCT AGACAGCCCC 1260
CGAGCCCTTC GGACAGAGGT CACCTCCAAG ATGTTCCAGA AGACGAGGAC TGGGTCTTTT 1320
CAGATCCGGC GGGGGGAAGT GGGTGCAGC CAGGTGCAGC TGCGGGCCCT TGAGCACGTG 1380
GATGGGACGC ACGTGTGCCA GCTGCCGGAG GACCAGAAGG GCAACATCCA TCTGAAACCT 1440
TCCTTCTCCG ACGGCCTCAA GATGGACGCG GGCATCATCT GTGATGTGTG CACCTGCGAG 1500
CTGCAAAAAG AGGTGCGGTC AGCTCGCTGC AGCTTCAACG GAGACTTCGT GTGCGGACAG 1560
TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT 1620
GACATTACAG CTTGCCCTGC GGAGGGCGAG GACAAGCCGT GCTCCGGCCG TGGGGAGTGC 1680
CAGTGCGGGC ACTGTGTGTG CTACGGCGAA GGGCGTACG AGGGTCAGTT CTGCGAGTAT 1740
GACAACCTTC AGTGTCCCCG CACTTCCGGG TTCCTCTGCA ATGACCGAGG ACGCTGCTCC 1800
ATGGGCCAGT GTGTGTGTGA GCCTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCCTCAGC 1860
AATGCCACCT GCATCGACAG CAATGGGGGC ATCTGTAATG GACGTGGCCA CTGTGAGTGT 1920
GGCCGCTGCC ACTGCCACCA GCAGTCGCTC TACACGGACA CCATCTGCGA GATCAACTAC 1980
TCGGCGATCC ACCCGGGCCT CTGCGAGGAG CTACGCTCCT GCGTGCAGTG CCAGGCGTGG 2040
GGCACCAGCG AGAAGAAGG GCGCAATGTA GAGGAATGCA ACTCAAGGT CAAGATGGTG 2100
GACGAGCTTA AGAGAGCCGA GGAGGTGGTG GTGCGCTGCT CTTCCGGGA CGAGGATGAC 2160
GACTGCACCT ACAGCTACAC CATGGAAGGT GACGGCGCCC CTGGGCCCAA CAGCACTGTC 2220
CTGGTGCACA AGAAGAAGGA CTGCCCTCCG GGCTCCTTCT GGTGGCTCAT CCCCCTGCTC 2280
CTCCTCCTCC TGCCGCTCCT GGGCCTGCTA CTGCTGCTAT GCTGGAAGTA CTGTGCCTGC 2340
TGCAAGGCCT GCCTGGCACT TCTCCCGTGC TGCAACCGAG GTCACATGGT GGGCTTTAAG 2400
GAAGACCACT ACATGCTGCG GGAGAACCTG ATGGCTCTG ACCACTTGA CACGCCCATG 2460
CTGCGCAGCG GGAACCTCAA GGGCCGTGAC GTGGTCCGCT GGAAGGTCAC CAACAACATG 2520
CAGCGGCCTG GCTTTGCCAC TCATGCCGCC AGCATCAACC CCACAGAGCT GGTGCCCTAC 2580
GGGCTGTCTT TGCGCCTGGC CCGCCTTTGC ACCGAGAACC TGCTGAAGCC TGACACTCGG 2640
GAGTGCGCCC AGCTGCGCCA GGAGGTGGAG GAGAACCTGA ACCAGGTCTA CAGGCAGATC 2700
TCCGGTGTAC ACAAGCTCCA GCAGACCAAG TTCCGGCAGC AGCCCAATGC CGGGAAGAA 2760
CAAGACCACA CCATTGTGGA CACAGTACAG ATGGGCCCC GCTCGGCCAA GCCGGCCCTG 2820
CTGAAGCTTA CAGAGAAGCA GGTGGAACAG AGGGCCTTCC ACGACCTCAA GGTGGCCCCC 2880
GGCTACTACA CCCTCACTGC AGACCAGGAC GCCCGGGGCA TGGTGGAGTT CCAGGAGGGC 2940
GTGGAGCTGG TGGACGTACG GGTGCCCTTC TTTATCCGGC CTGAGGATGA CGACGAGAAG 3000
CAGCTGCTGG TGGAGGCCAT CGACGTGCCC GCAGGCACTG CCACCTCGG CCGCCGCTG 3060

```



5 GTAAACATCA CCATCATCAA GGAGCAAGCC AGAGACGTGG TGTCTTTTGA GCAGCCTGAG 3120  
TTCTCGGTCA GCCGCGGGGA CCAGGTGGCC CGCATCCCTG TCATCCGGCG TGTCTGGAC 3180  
GGCGGGAAGT CCCAGGTCTC CTACCGCACA CAGGATGGCA CCGCGCAGGG CAACCGGGAC 3240  
TACATCCCCG TGGAGGGTGA GCTGCTGTTC CAGCCTGGGG AGGCCTGGAA AGAGCTGCAG 3300  
GTGAAGCTCC TGGAGCTGCA AGAAGTTGAC TCCCTCCTGC GGGGCCGCCA GGTCCGCCGT 3360  
TTCCACGTCC AGCTCAGCAA CCCTAAGTTT GGGGCCACC TGGGCCAGCC CCACTCCACC 3420  
ACCATCATCA TCAGGGACCC AGATGAAGT GACCGGAGCT TCACGAGTCA GATGTTGTCA 3480  
TCACAGCCAC CCCCTCACGG CGACCTGGGC GCCCGCAGA ACCCAATGC TAAGGCCGCT 3540  
GGGTCCAGGA AGATCCATT CAACTGGCTG CCCCTTCTG GCAAGCCAAT GGGGTACAGG 3600  
10 GTAAAGTACT GGATTCAGGG TGAATCCGAA TCCGAAGCCC ACCTGCTCGA CAGCAAGGTG 3660  
CCCTCAGTGG AGCTCACCAA CCTGTACCCG TATTGCGACT ATGAGATGAA GGTGTGCGCC 3720  
TACGGGGCTC AGGGCGAGGG ACCCTACAGC TCCCTGCTGT CCTGCCGCAC CCACCAGGAA 3780  
GTGCCCAGCG AGCCAGGGCG TCTGGCCTTC AATGTCGTCT CCTCCACGGT GACCCAGCTG 3840  
AGCTGGGCTG AGCCGGCTGA GACCAACGGT GAGATCACAG CCTACGAGGT CTGCTATGGC 3900  
15 CTGGTCAACG ATGACAACCG ACCTATTGGG CCCATGAAGA AAGTGCTGGT TGACAACCCT 3960  
AAGAACCGBA TGCTGCTTAT TGAGAACCTT CGGGAGTCCC AGCCCTACCG CTACACGGTG 4020  
AAGGCGCGCA ACGGGGCCGG CTGGGGGCCT GAGCGGGAGG CCATCATCAA CCTGGCCACC 4080  
CAGCCCAAGA GGCCCATGTC CATCCCCATC ATCCCTGACA TCCCTATCGT GGACGCCCAG 4140  
AGCGGGGAGG ACTACGACAG CTTCTTATG TACAGCGATG ACGTTCTACG CTCTCCATCG 4200  
20 GGCAGCCAGA GGCCAGCGT CTCCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGGAC 4260  
TTTGCTTCC CGGGCAGCAC CAACTCCCTG CACAGGATGA CCACGACCAG TGCTGCTGCC 4320  
TATGGCACCC ACCTGAGCCC ACACGTGCCC CACCGCTGTC TAAGCACATC CTCCACCTC 4380  
ACACGGGACT ACAACTCACT GACCCGCTCA GAACACTCAC ACTCGACCAC ACTGCCGAGG 4440  
GACTACTCCA CCCTCACCTC CGTCTCCTCC CACGACTCTC GCCTGACTGC TGGTGTGCCC 4500  
25 GACACGCCCA CCCGCTGGT GTTCTCTGCC CTGGGGCCCA CATCTCTCAG AGTGAGCTGG 4560  
CAGGAGCCGC GGTGCGAGCG GCCGCTGCAG GGCTACAGTG TGGAGTACCA GCTGCTGAAC 4620  
GGCGGTGAGC TGCATCGGCT CAACATCCCC AACCTGCCC AGACCTCGGT GGTGGTGGAA 4680  
GACCTCCTGC CCAACCACTC CTACGTGTTT CGCGTGCGGG CCCAGAGCCA GGAAGGCTGG 4740  
30 GGCCGAGAGC GTGAGGGTGT CATCACCATT GAATCCCAGG TGCACCCGCA GAGCCCACTG 4800  
TGTCCCCTGC CAGGCTCCGC CTTCACTTTG AGCACTCCCA GTGCCCCAGG CCCGCTGGTG 4860  
TTCATGCCCC TGAGCCCAGA CTCGCTGCAG CTGAGCTGGG AGCGGCCACG GAGGCCCAAT 4920  
GGGGATATCG TCGGCTACCT GGTGACCTGT GAGATGGCCC AAGGAGGAGG GCCAGCCACC 4980  
GCATTCGGG TGGATGGAGA CAGCCCCGAG AGCCGGCTGA CCGTGCCGGG CCTCAGCGAG 5040  
35 AACGTGCCCT ACAAGTTCAA GGTGCAAGCC AGGCACTG AGGGCTTCGG GCCAGAGCGC 5100  
GAGGGCATCA TCACCATAGA GTCCAGGAT GGAGGACCTT TCCCGCAGCT GGGCAGCCGT 5160  
GCCGGGCTCT TCCAGCACCC GCTGCAAAGC GAGTACAGCA GCATCACACC CACCCACACC 5220  
AGCGCCACCG AGCCCTTCTT AGTGGATGGG CCGACCTGG GGGCCAGCA CCTGGAGGCA 5280  
GGCGGCTCCC TCACCCGGCA TGTGACCCAG GAGTTTGTGA GCCCGACACT GACCACCAGC 5340  
40 GGAACCCTTA GCACCCACAT GGACCAACAG TTCTTCCAAA CTTGACCGCA CCCTGCCCA 5400  
CCCCCGCCAT GTCCCACTAG GCTCCCTCCC GACTCCTCTC CCGGAGCCTC CTCAGCTACT 5460  
CCATCCTTGC ACCCTGGGG GCCCAGCCCA CCCGCATGCA CAGAGCAGGG GCTAGGTGTC 5520  
TCCTGGGAGG CATGAAGGGG GCAAGGTCCG TCCTCTGTGG GCCCAAACCT ATTTGTAACC 5580  
AAAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTCTG CACTTAATAA ATGGTTTTCG 5640  
TACTG

Seq ID NO: 561 Protein sequence  
Protein Accession #: NP\_000204.1

50 1 11 21 31 41 51  
MAGPRPSPWA RLLLAALISV SLSGTLANRC KKAPVKSCTE CVRVDKDCAY CTDEMFRDRR 60  
CNTQAEELAA GCQRESIVVM ESSFQITEET QIDTTLRRSQ MSPQGLRVRL RPGEERHFEL 120  
EVFEPLSEPV DLYILMDFSN SMSDDLNLK KMGQNLARVL SGLTSDYTIG FGKFDKVSFV 180  
55 PQTDMRPEKL KEPWPNSDPP FSFKNVISLT EDVDEFNRKL QGERISGNLD APEGGFDAIL 240  
QTAVCTRDIG WRPDSTHLLV FSTESAFHYE ADGANVLAGI MSRNDERCHL DTTGTYTQYR 300  
TQDYPSVPTL VRLAKHNI I PIFAVTNYSY SYEKLHTYF PVSSLGVLQE DSSNIVELLE 360  
EAFNRIRSNL DIRALDSPRG LRTEVTSKMF QKTRTGSFHI RRGEVGIYQV QLRALHVDG 420  
THVCQLPEDQ KGNHLKPSF SDGLKMDAGI ICDVCTCELQ KEVRSARCSF NGDFVCGQCV 480  
60 CSEGWSGQTC NCSTGSLSDI QPCLRGECQC PCSGRGECQC GHCVCYGEGR YEQGFCEYDN 540  
FQCPRTSGLF CNDRGRCMGM QCVCEPGWTG PSCDCPLSNA TCIDSNGGIC NGRGHCECGR 600  
CHCHQQSLYT DTICEINYSI IHPGLCEDLR SCVQCQAWGT GEKKGRTCEE CNFKVKMVDE 660  
LKRAEEVVVR CSFRDEDDDC TYSYTMEDGD APGPNSTVLV HKKDCPPGS FWLIPLLL 720  
LLPLALLLLL LCWKYCACCK ACLALLPCCN RGHMVGFKED HYMLRENLM A SDHLDTPMLR 780  
SGNLKGRDVV RWKVTNNMQR PGFATHAASI NPTELVPYGL SLRLARLCTE NLLKPDTRC 840  
65 AQLRQEVEEN LNEVYRQISG VHKLQQTKEF QQPNAKKQD HTIVDTVIMA PRSAKPALLK 900  
LTEKQVEQRA FHDLKVAPGY YTLTAQDAR GMVEFQEGVE LVDVRVPLFI RPEDDDEKQL 960  
LVEAIDVPAG TATLGRRLVN ITIIEQARD VVSFEQPEFS VSRGDQVARI PVIRRVLDGG 1020  
KSQVSYRTQD GTAQGNRDYI PVEGELLFQP GEAWKELQVK LLELQEVDSL LRGRQVRRFH 1080  
70 VQLSNPKFGA HLGQPHSTTI IIRDPDELDR SFTSQMLSSQ PPPHGDLAGP QNPNAKAAGS 1140  
RKIHFNWLPF SGKPMGYRVK YWIQGDSESE AHLDSKVPV VELTNLYPYC DYEMKVCAYG 1200  
AQGEQPYSSL VSCRTHQEVV SEPGRILAFNV VSSVTQLSW AEPATNGEI TAYEVCYGLV 1260  
NDDNRPIGPM KKVLDNPKN RMLLIENLRE SQPYRYTVKA RAGAGWGPFR EAIINLATQP 1320  
KRPMIPIIP DIPIVDAQSG EDYDSFLMYS DDVLRSPSGS QRPSVSDDE HLVNGRMDFA 1380  
75 FPGSTNSLHR MTTTSAAYG THLSFPHVPHR VLSTSTLSTR DYNLSTRSEH SHSTTLPRDY 1440  
STLTSVSSH SRLTAGVPDT PTRLVFSALG PTLRVSWQE PRCEPLQGY SVEYQLLNGG 1500  
ELHRLNIPNP AQTSVVVDL LPNHSYVFRV RAQSQEGWGR EREGVITIES QVHPQSPLCP 1560  
LPGSAFTLST PSAPGPLVFT ALSPDSLQLS WERPRRPNGD IVGYLVTCM AQGGGPATAF 1620  
RVDGDSPEER LTVPGLENSV PYKFKVQART TEGFQPEREG IITIESQDGG PFPQLGSRAG 1680  
80 LFQHPQLQSEY SSITTTHTSA TEPFLVDGPT LGAQHLEAGG SLTRHVTQEF VSRTLTTSST 1740  
LSTHMDQQFF QT

Seq ID NO: 562 DNA sequence  
Nucleic Acid Accession #: NM\_013332.1  
Coding sequence: 1..63

85 1 11 21 31 41 51  
| | | | | |



GCACGAGGGC GCTTTTGTCT CCGGTGAGTT TTGTGGCGGG AAGCTTCTGC GCTGGTGCTT 60  
AGTAACCGAC TTTCCTCCGG ACTCCTGCAC GACCTGCTCC TACAGCCGGC GATCCACTCC 120  
CGGCTGTTC CCGGAGGGT CCAGAGGCCT TTCAGAAGGA GAAGGCAGCT CTGTTTCTCT 180  
GCAGAGGAGT AGGGTCCTTT CAGCCATGAA GCATGTGTG AACCTCTACC TGTTAGGTGT 240  
GGTACTGACC CTACTCTCCA TCTTCGTTAG AGTGATGGAG TCCCTAGAAG GCTTACTAGA 300  
GAGCCCATCG CCTGGGACCT CCTGGACCAC CAGAAGCCAA CTAGCCAACA CAGAGCCCAC 360  
CAAGGGCCTT CCAGACCATC CATCCAGAAG CATGTGATAA GACCTCCTTC CATACTGGCC 420  
ATATTTTGA ACATGACCT AGACATGTCC AGATGGGAGT CCCATTCTTA GCAGACAAGC 480  
TGAGCACCGT TGTAACCAGA GAACTATTAC TAGGCCTTGA AGAACCTGTC TAACTGGATG 540  
CTCATTGCCT GGGCAAGGCC TGTTTAGGCC GGTGCGGTG GCTCATGCCT GTAATCCTAG 600  
CACTTTGGGA GGCTGAGGTG GGTGGATCAC CTGAGGTCAG GAGTTCGAGA CCAGCCTCGC 660  
CAACATGGCG AAACCCCATC TCTACTAAAA ATACAAAAGT TAGCTGGGTG TGGTGGCAGA 720  
GGCCTGTAAT CCCAGTTCCT TGGGAGGCTG AGGCGGGAGA ATTGCTTGAA CCCGGGGACG 780  
GAGGTTGCAG TGAACCGAGA TCGCACTGCT GTACCCAGCC TGGGCCACAG TGCAAGACTC 840  
CATCTCAAAA AAAAAAAGAA AAGAAAAAGT CTGTTTAAATG CACAGGTGTG AGTGGATTGC 900  
TTATGGCTAT GAGATAGGTT GATCTCGCCC TTACCCCGGG GTCTGGTGTA TGCTGTGCTT 960  
TCCTCAGCAG TATGGCTCTG ACATCTCTTA GATGTCCCAA CTTCAGCTGT TGGGAGATGG 1020  
TGATATTTTC AACCTACTT CCTAAACATC TGTCTGGGGT TCCTTTAGTC TTGAATGTCT 1080  
TATGCTCAAT TATTTGGTGT TGAGCCTCTC TTCCACAAGA GCTCCTCCAT GTTTGGATAG 1140  
CAGTTGAAGA GGTGTGTGG GTGGCTGTT GGGAGTGAGG ATGGAGTGT CAGTGCCCAT 1200  
TTCTCATTTT ACATTTTAAA GTCGTTCTCTC CAACATAGTG TGTATTGGTC TGAAGGGGGT 1260  
GGTGGGATGC CAAAGCCTGC TCAAGTTATG GACATTGTGG CCACCATGTG GCTTAAATGA 1320  
TTTTTTCTAA CTAATAAAGT GGAATATATA TTTCAAAAA AAAAAAAAAA AA

Seq ID NO: 563 Protein sequence  
Protein Accession #: NP\_037464.1

1 11 21 31 41 51  
MKHVLNLYLL GVVLTLLSIF VRVMESLEGL LESPSPGTSW TTRSQLANTE PTKGLPDHPS 60  
RSM

Seq ID NO: 564 DNA sequence  
Nucleic Acid Accession #: NM\_023915.1  
Coding sequence: 250..1326

1 11 21 31 41 51  
GGCAGGAGGG TTTCGTTTTC ATGCTTTACC AGAAAATCCA CTTCCCTGCC GACCTTAGTT 60  
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120  
GTGAATGGAC AGCCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180  
CCCACGCCCT AATCGTCCCC AAGTGTTCCT TGACACGCAT CTTTGCTTAC AGTGCATCAC 240  
AACTGAAGAA TGGGGTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCACGGC 300  
CAAGAGAGTC ACAATTCAGG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCCTTCAC 360  
AATGAATTTG ACACAATTGT CTTGCCGGTG CTTTATCTCA TTATATTTGT GGCAAGCATC 420  
TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAAC CAGCTTCATA 480  
TTCTATCTCA AAAACATAGT GGTTCAGAC CTCATAATGA CGCTGACATT TCCATTTCTGA 540  
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600  
TCAGTTTTGT TTTATGCAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCAT 660  
GATCGCTATC TGAAGGTGGT CAAGCCATTT GGGGACTCTC GGATGTACAG CATAACCTTC 720  
ACGAAGGTTT TATCTGTTTG TGTGTTGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780  
ATCCTGACAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAAA ACTTAAAAGT 840  
CCTTTGGGGG TCAAAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900  
GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960  
AGGCAATTCA TAAGTCAGTC AAGCCGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT 1020  
GTGGCTGTGT TTTTACCTG CTTTCTACCA TATCACTTGT GCAGAATTCC TTTTACTTTT 1080  
AGTCACTTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCTATATTA CTGCAAAGAA 1140  
ATTACACTTT TCTGTCTGC GTGTAATGTT TGCTTGGATC CAATAATTTA CTTTTTCATG 1200  
TGTAGGTCAT TTTCAAGAAG GCTGTTCAAA AAATCAAATA TCAGAACCAG GAGTGAAAGC 1260  
ATCAGATCAC TGCAAAGTGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTACTACTGAT 1320  
GTGTAGGCCT TTTATTGTTT GTTGAATCG ATATGTACAA AGTGTAATAA AATGTTTCTT 1380  
TTCATTATCC TTAATAAAAA AA

Seq ID NO: 565 Protein sequence  
Protein Accession #: NP\_076404

1 11 21 31 41 51  
MGFNLTLAKL PNNEHLHGQES HNSGNRSDGP GKNTTLHNEF DTIVLPVLYL IIFVASILLN 60  
GLAVWIFFHI RNKTSFIFYL KNIVVADLIM TLTFPFRIVH DAGFGPWYFK FILCRYTSVL 120  
FYANMYTSIV FLGLISIDRY LKVVKPFGLD RMYSTFTKV LSVCVWVIMA VLSLPNIILT 180  
NGQPTEDNIH DCSKLSPLG VKWHTAVTYV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQF 240  
ISQSSRKRGH NQSIRVVAV FFTCFLPYHL CRIPFTFSLH DRLLDESAQK ILYYCKEITL 300  
FLSACNVCLD PIIFYFMCRS FSRRFLFKSN ITRSESIRS LQSVRRSEVR IYYDYTDV

Seq ID NO: 566 DNA sequence  
Nucleic Acid Accession #: NM\_005365.1  
Coding sequence: 1..948

1 11 21 31 41 51  
ATGTCTCTCG AGCAGAGGAG TCCGCACTGC AAGCCTGATG AAGACCTTGA AGCCCAAGGA 60  
GAGCACTTGG GCCTGATGGG TGCACAGGAA CCCACAGGCG AGGAGGAGGA GACTACCTCC 120  
TCCCTCTGACA GCAAGGAGGA GGAGGTGTCT GCTGCTGGGT CATCAAGTCC TCCCCAGAGT 180  
CCTCAGGGAG GCGCTTCCTC CTCCATTTCC GTCTACTACA CTTTATGGAG CCAATTCGAT 240  
GAGGGCTCCA GCAGTCAAGA AGAGGAAGAG CCAAGCTCCT CGGTGACCCC AGCTCAGCTG 300  
GAGTTCATGT TCCAAGAAGC ACTGAAATTG AAGGTGGCTG AGTTGGTTCA TTTCTGCTC 360



CACAAATATC GAGTCAAGGA GCCGGTCACA AAGGCAGAAA TGCTGGAGAG CGTCATCAAA 420  
AATTACAAGC GCTACTTTCC TGTGATCTTC GGCAAAGCCT CCGAGTTCAT GCAGGTGATC 480  
TTTGGCACTG ATGTGAAGGA GGTGGACCCC GCCGGCCACT CCTACATCCT TGTCAGTCTG 540  
CTTGGCCTCT CGTGGCAGAT CATGCTGGGT GATGGTCATA GCATGCCCAA GGCCGCCCTC 600  
5 CTGATCATTTG TCCTGGGTGT GATCCTAACC AAAGACAAC TCGCCCCCTGA AGAGGTTATC 660  
TGGGAAGCGT TGAGTGTGAT GGGGGTGTAT GTTGGGAAGG AGCACATGTT CTACGGGGAG 720  
CCCAGGAAGC TGCTCACCCA AGATTGGGTG CAGGAAAAC TACCTGGAGTA CCGGCAGGTG 780  
CCCGGCAGTG ATCCTGCGCA CTACGAGTTC CTGTGGGGTT CCAAGGCCCA CGCTGAAACC 840  
10 AGCTATGAGA AGGTCATAAA TTATTTGGTC ATGCTCAATG CAAGAGAGCC CATCTGCTAC 900  
CCATCCCTTT ATGAAGAGGT TTTGGGAGAG GAGCAAGAGG GAGTCTGA

Seq ID NO: 567 Protein sequence  
Protein Accession #: NP\_005356.1

1 11 21 31 41 51  
| | | | | |  
MSLEQRSPHC KPDEDLAQG EDLGLMGAQE PTGEEETTS SSDSKEEEVS AAGSSSPQPS 60  
PQGGASSSIS VYVTLWSQFD EGSSSQEHEE PSSSVDPQL EFMFQELKL KVAELVHFL 120  
20 HKYRVKEPVT KAEMLESVIK NYKRYFPVIF GKASEFMQVI FGTDVKEVDP AGHSYILVTA 180  
LGLSCDSMLG DGHSMFKAAL LIIVLGIVLT KDNCAPPEVI WEALSVMGVY VGKEHMFYGE 240  
PRKLLTQDWV QENYLEYRQV PGSDPAHYEF LWGSKAHAET SYEKVINYL MLNAREPICY 300  
PSLYEEVLGE EQEGV

Seq ID NO: 568 DNA sequence  
Nucleic Acid Accession #: NM\_014400  
Coding sequence: 86..1126

1 11 21 31 41 51  
| | | | | |  
30 GGTACTCAT CCTGGGCTCA GGTAAAGAGG CCCGAGCTCG GAGGCGGCAC ACCCAGGGGG 60  
GACGCCAAGG GAGCAGGACG GAGCCATGGA CCCCGCCAGG AAAGCAGGTG CCCAGGCCAT 120  
GATCTGGACT GCAGGCTGGC TGCTGCTGCT GCTGCTTCGC GGAGGAGCGC AGGCCCTGGA 180  
GTGCTACAGC TCGGTGCAGA AAGCAGATGA CGGATGCTCC CCGAACAAGA TGAAGACAGT 240  
35 GAAGTGCGCG CCGGGCGTGG ACGTCTGCAC CGAGGCCGTG GGGGCGGTGG AGACCATCCA 300  
CGGACAATTC TCGCTGGCAG TGCSGGGTG CCGTTCGGGA CTCCCCGGCA AGAATGACCG 360  
CGGCCTGGAT CTTACGGGC TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420  
CTGCAACGCC AAGTCAACC TCACCTCGCG GCGCTCGAC CCGGCAGGTA ATGAGAGTGC 480  
ATACCCGCCC AACGGCGTGG AGTGCTACAG CTGTGTGGG CTGAGCCGGG AGGCGTGCCA 540  
40 GGTACATCG CCGCCGTCG TGAGCTGCTA CAACGCCAGC GATCATGTCT ACAAGGGCTG 600  
CTTCGACGGC AACGTACCT TGACGGCAGC TAATGTGACT GTGCTCTGCT CTGTCCGGGG 660  
CTGTGTCCAG GATGAATCT GCACTCGGGA TGGAGTAACA GGCCCAAGGT TCACGCTCAG 720  
TGGCTCCTGT TGCCAGGGGT CCGCTGTAA CTCTGACCTC CGCAACAAGA CCTACTTCTC 780  
CCCTCGAATC CCACCCCTG TCCGGCTGCC CCCTCCAGAG CCCACGACTG TGGCCTCAAC 840  
45 CACATCTGTC ACCACTTCTA CCTCGGCCCC AGTGAGACCC ACATCCACCA CCAAACCCAT 900  
GCCAGCGCCA ACCAGTCAGA CTCCGAGACA GCGAGTAGAA CACGAGGCCCT CCCGGGATGA 960  
GGAGCCCAGG TTGACTGGAG GCGCCGCTGG CCACCAGGAC CGCAGCAATT CAGGGCAGTA 1020  
TCCTGCAAAA GGGGGGCCCC AGCAGCCCCA TAATAAAGGC TGTGTGGCTC CCACAGCTGG 1080  
ATTGGCAGCC CTTCTGTTGG CCGTGGCTGC TGGTGTCTTA CTGTGAGCTT CTCCACCTGG 1140  
50 AAATTTCCCT CTCACCTACT TCTCTGGCCC TGGGTACCCC TCTTCTCATC ACTTCTGTT 1200  
CCCACCACTG GACTGGGCTG GCGGAGCCCC TGTTTTTCCA ACATTCCTCA GTATCCCCAG 1260  
CTTCTGCTGC GCTGGTTTGC GGCTTTGGGA AATAAAATAC CGTTGTATAT ATTCTGGCAG 1320  
GGGTGTTCTA GCTTTTGGAG GACAGCTCCT GTATCCTTCT CATCCTTGTG TCTCCGCTTG 1380  
TCCTCTTGTG ATGTTAGGAC AGAGTGAAGC AAGTCAGCTG TCACGGGGAA GGTGAGAGAG 1440  
55 AGGATGCTAA GCTTCTACT CACTTTCTCC TAGCCAGCCT GGACTTTGGA GCGTGGGGTG 1500  
GGTGGGACAA TGGCTCCCCA CTCTAAGCAC TGCCTCCCT ACTCCCCGCA TCTTTGGGGA 1560  
ATCGGTTCCC CATATGTCTT CCTTACTAGA CTGTGAGCTC CTCGAGGGCA GGGACCGTGC 1620  
CTTATGTCTG TGTGTGATCA GTTTCTGGCA CATAAATGCC TCAATAAAGA TTTAATTACT 1680  
TTGTATAGTG AAAAAAAA

Seq ID NO: 569 Protein sequence  
Protein Accession #: NP\_055215

1 11 21 31 41 51  
| | | | | |  
65 MDPARKAGAQ AMIWTAGWLL LLLLRRGAQA LECYSCVQKA DDGCSFNKMK TVKCAPGVDV 60  
CTEAVGAVET IHGQFSLAVX GCGSGLPGKN DRGLDLHGLL AFIQLQCCAQ DRCNAKLNLT 120  
SRALDPAGNE SAYPPNGVEC YSCVGLSREA CQGTSPPVVS CYNASDHVYK GCFDGNVTLT 180  
AANVTVSLPV RGCVDDEFCT RDGVTGPGFT LSGSCCQSR CNSDLRNKTY FSPRIPLVR 240  
70 LPPPEPTTVA STTSVTSTT APVRPTSTT PMPAPTSQTP RQGEHEASR DEEPRLTGGA 300  
AGHQDRNSG QYPAKGGPQQ PHNKGCVAPT AGLAALLLAV AAGVLL

Seq ID NO: 570 DNA sequence  
Nucleic Acid Accession #: NM\_005329.1  
Coding sequence: 1..1662

1 11 21 31 41 51  
| | | | | |  
75 ATGCCGCTGC AGCTGACGAC AGCCCTGCGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60  
GTGCTGGGTG GCATCTGGC AGCCTATGTG ACGGGCTACC AGTTCATCCA CACGAAAAG 120  
80 CACTACCTGT CCTTCGGCCT GTACGGCGCC ATCCTGGGCC TGCACCTGCT CATTGAGAGC 180  
CTTTTTCCTT TCCTGGAGCA CCGGCGCATG CGACGTGCCG GCCAGGCCCT GAAGCTGCCC 240  
TCCCCGCGGC GGGGCTCGGT GGCACCTGTC ATTGCCGCGT ACCAGGAGGA CCCTGACTAC 300  
TTGCGCAAGT GCCTGCGCTC GGCCAGCGC ATCTCCTTCC CTGACCTCAA GGTGGTCATG 360  
85 GTGTGGATG GCAACCGCCA GGAGGAGGCC TACATCTGG ACATCTTCCA CGAGGTGCTG 420  
GGCGGCACCG AGCAGGCCCG CTCTTTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG 480  
GGTGAGACGG AGGCCAGCCT GCAGGAGGGC ATGGACCGTG TCGGGGATGT GGTGCGGGCC 540  
AGCACCTTCT CGTGATCAT GCAGAAGTGG GGAGGCAAGC GCGAGGTCAT GTACACGGCC 600



TTCAAGGCC TCGGCGATT GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660  
 GATCCAGCCT GCACCATCGA GATGCTTCGA GTCCCTGGAGG AGGATCCCCA AGTAGGGGGA 720  
 GTCCGGGGAG ATGTCCAGAT CCTCAACAAG TACGACTCAT GGATTTCCTT CCTGAGCAGC 780  
 GTGCGGTACT GGATGGCCTT CAACGTGGAG CGGGCCTGCC AGTCCTACTT TGGCTGTGTG 840  
 CAGTGTATTA GTGGGCCCTT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCCTGGAG 900  
 GACTGGTACC ATCAGAAAGT CCTAGGCAGC AAGTGCAGCT TCGGGGATGA CCGGCACCTC 960  
 ACCAACCCGAG TCCTGAGCCT TGGTACCGA ACTAAGTATA CCGCGCGCTC CAAGTGCCTC 1020  
 ACAGAGACCC CCACTAAGTA CCTCCGGTGG CTCAACCAGC AAACCCGCTG GAGCAAGTCT 1080  
 TACTTCCGGG AGTGGCTCTA CAACTGCTGT TGGTTCCTA AGCACCACCT CTGGATGACC 1140  
 TACGAGTCAG TGGTCACGGG TTTCTTCCCC TTCTTCCTCA TTGCCACGGT TATACAGCTT 1200  
 TTCTACCGGG GCCGCATCTG GAACATTCTC CTCTTCCTGC TGACGGTGCA GCTGGTGGGC 1260  
 ATTATCAAGG CCACCTACGC CTGCTTCCTT CGGGGCAATG CAGAGATGAT CTTTCATGTCC 1320  
 CTCTACTCCC TCCTCTATAT GTCCAGCCTT CTGCCGGCCA AGATCTTTGC CATTGCTACC 1380  
 ATCAACAAAT CTGGCTGGGG CACCTCTGGC CGAAAAACCA TTGTGGTGAA CTTTCATTGGC 1440  
 CTCATTCCCTG TGTCCATCTG GGTGGCAGTT CTCTTGGGAG GGCTGGCCTA CACAGCTTAT 1500  
 TGCCAGGACC TGTTCAGTGA GACAGAGCTA GCCTTCCTTG TCTCTGGGGC TATACTGTAT 1560  
 GGCTGCTACT GGGTGGCCCT CCTCATGCTA TATCTGGCCA TCATCGCCCC GCGATGTGGG 1620  
 AAGAAGCCGG AGCAGTACAG CTTGGCTTTT GCTGAGGTGT GA

Seq ID NO: 571 Protein sequence  
 Protein Accession #: NP\_005320.1

1 11 21 31 41 51  
 | | | | |  
 MPVQLTTLR VVGTSLFALA VLGGILAAYV TGYQFIHTEK HYLSFGLYGA ILGLHLLIQS 60  
 LFAFLEHRRM RRAGQALKLP SPRRGSVALL IAAYQEDPDY LRKCLRSAQR ISFPDLKVVM 120  
 VVDGNRQEDA YMLDIFHEVL GGTEQAGFFV WRSNFHEAGE GETEASLQEG MDRVRDVVRA 180  
 STFSCIMQKW GSKREVMTA FKALGDSVDY IQVCDSDTVL DPACTIEMLR VLEEDPQVGG 240  
 VGGDVQILNK YDSWISFLSS VRYWMAFNVE RACQSYFGCV QCISGPLGMY RNSLLQQFLE 300  
 DWYHQKFLGS KCSFGDDRHL TNRVLSLGYR TKYTARKSKL TETPTKYLRW LNQQTRWSKS 360  
 YFREWLNSL WPHKHHLWMT YESVVTGFFP FFLIATVIQL FYRGRIWNIL LFLLTQVLVG 420  
 IIKATYACFL RGNAEMIFMS LYSLLYMSL LPAKIFAIAT INKSGWGTSG RKTIVVNFIF 480  
 LIPVSIWVAV LLGGLAYTAY CQDLFSETEL AFLVSGAILY GCYWVALLML YLAIIRRCG 540  
 KKPEQYSLAF AEV

Seq ID NO: 572 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 148-7095

1 11 21 31 41 51  
 | | | | |  
 CACACATACG CACGCACGAT CTCACCTTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAAC ATTTCCCTTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCCT CGCTTGCAAT 180  
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTTGTTGAAG AGATTGGCTG GTCCATATCA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG 300  
 AAATATCCAA CATGTAATAG CCCAAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA 360  
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA 420  
 AACACATTCA TTCATAACAC TGGGAAAACA GTGGAAATTA ATCTCACTAA TGAATACCGT 480  
 GTCAGCGGAG GAGTTTCAGA AATGGTGTGT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 AAATGCAATA TGTCATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600  
 GAGATGCAAA TCTACTGCTT TGATGCGGAC GTATTTTCAA GTTTTGAGGA AGCAGTCAAA 660  
 GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAAATTG 720  
 GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTC GTTTTGGGAA GCAGGCTGCT 780  
 TTAGATCCAT TCATACTGTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTTAC 840  
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTGTGTG AAGTTCTTAC AATGCAACAA 960  
 TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTGAGAGCA ACAGTACAAG 1020  
 TTCTCTAGAC AGGTGTTTTT CTCTACACTG GGAAGGGAAG AGATTCATGA AGCAGTTTGT 1080  
 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA 1140  
 TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCACT TTTGTACCAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380  
 AATCCTGAAC TTGATCTTTT CCCTGACTTA ATTGGAACTG AAGAAATAAT CAAGGAGGAG 1440  
 GAAGAGGGA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500  
 AACCAATCA GGAAGGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560  
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCAACAA GAGGAAGTGA ATCTCTGGA 1620  
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCAAGTCA TAAATTAGCC 1680  
 ACAGAAAAAG ATATTTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAACTG TTCTTAGATC TCCACATATG 1800  
 AACTTGTCCG GGAATGAGA ATCTTAAAT CAGTTTCTTA TAACAGAATA TGAGGAGGAG 1860  
 AGTTTATTGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920  
 GCAACTTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980  
 GAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040  
 GAAGATTCAA CTTTCATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAT 2100  
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCC ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTTCTCC AGACTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220  
 TCCTTTTCTG CAGGCCAGT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA 2280  
 CATTATTCTA CTTTGCCTA CTTCCCACT GAGGTAACAC CTCATGCTTT TACCCATCC 2340  
 TCCAGACAAC AGGATTTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCACCG 2400  
 GTATACAATG GTGAGACACC TCTTCAACCT TCCTACAGTA GTGAAGTCTT TCCTCTAGTC 2460  
 ACCCCTTTGT TGCTTGACAA TCAGATCTCT CAGACTACCC CTGCTGCTTC AAGTAGTGAT 2520  
 TCGGCCTTGC ATGCTACGCC TGTATTTCCC AGTGTGATG TGTCAATTTA ATCCATCCTG 2580  
 TCTTCTATG ATGGTGCACC TTTGCTTCCA TTTTCTCTG CTTCTTCAG TAGTGAATTG 2640  
 TTTCGCCATC TGCATACAGT TTCTCAAATC CTTCCACAAG TTAATTCAGC TACCGAGAGT 2700  
 GATAAGGTGC CCTTGCATGC TTCTCTGCCA GTGGCTGGGG GTGATTTGCT ATTAGAGCCC 2760



	AGCCTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGGTAGTG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACCACCC	2880
	AGCAGTGATG	CCATGATGCA	TGCACGTTCT	TCAGGGCCTG	AACCTTCTTA	TGCCCTGTCT	2940
5	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTTTCTTACA	GTTCTGCAAT	ACCTGTGCAT	3000
	GATTCTGTGG	GTGTAACCTA	TCAGGGTTCC	TTATTTAGCG	GCCCTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCACCT	GTTTCTGTAG	CTGAATTTAC	ATATACAACA	3240
10	TCTGTGTTTG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGGAAATGAG	3300
	ACTGAACTGC	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGGCAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACTTTTTCA	TTCAACCTAC	ACATACTGTC	3540
15	TCTCAAGCAT	CTGGTGACAC	TTGCTTTAAA	CCTGTGCTTA	GTGCAAATC	AGAGCCAGCA	3600
	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATTG	CTACAACCTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAAACCTG	TCTTCCAGCT	GTGCCCAGTG	ATCCAATATT	GGTTGAAACC	3780
	CCCAAAGTTG	ATAAAATTAG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
20	AGTGAAAAAC	TGCTGCACTC	TACATCTGTA	CCAGTTTTTG	ATGTGTCGCC	TACTTCTCAT	3900
	ATGCACTCTG	CTTCACTTCA	AGGTTTGACC	ATTTCTTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTTTTGTATA	AAAGTGAAAG	TTCCCAACAA	GTGAGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCTAA	CGGCCAATTT	GGAGATTAAC	CAGGCCCATC	CCCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATTGATGAA	CCATTAAATA	CACTAATAAA	TAAGCTTATA	4140
25	CATTCCGATG	AAATTTTAA	CTCCACCAAA	AGTTCTGTGA	CTGGTAAGGT	ATTTGCTGGT	4200
	ATTCACACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTCTGTTCC	TATAGGAAAT	4260
	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCTGTAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
30	AGAGGTAGTG	ATGGCTTATC	CATTCTAAG	TGTATGTCAT	GCTCATCCTA	TAGAGAATCA	4500
	CAGGAAAAGG	TAATGAATGA	TTCAGACACC	CACGAAAACA	GTCTTATGGA	TCAGAATAAT	4560
	CCAAATCTCAT	ACTCACTATC	TGAGAATTCT	GAAGAAGATA	ATAGAGTCAC	AAGTGTATCC	4620
	TCAGACAGTC	AAACTGGTAT	GGACAGAAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACAATGATGG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
35	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCCTGACA	GTGATGAAGA	AAGTGGATCA	4800
	GGGCAAGGTA	CCTCAGATAG	CCTTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTTGCA	4860
	GACACTAATG	AAAAAGATGC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AATAACTCCT	4920
	GGATTCCAC	AGTCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAAGT	GTTCCACGTT	4980
	TCAGAGGCAG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	5040
40	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	5100
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	5220
	ATTTCAGATG	ATGTCGGAGC	AATTCCTAAT	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	5340
45	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	GCAGACAGCT	CCAACCACCC	AGACAACAAG	5400
	CACAAGAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460
	CTTGCTGAAA	AGGATGGCAA	ACTGATGTAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	5520
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTT	5580
	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	5640
50	AAAGGAAGGA	GAAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	5700
	TTTCTGGTCA	CTCAGAAGAG	TGTGCAAGTG	CTTGCCATT	ATACTGTGAG	GAATTTTACT	5760
	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	5820
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	5880
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTCTGC	5940
55	CACGTCAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	6000
	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAACACAT	CCGTTCAACA	6060
	AGAAATTAAT	TGGTACAAAC	TGAGGAGCAA	TATGTCTTCA	TTTATGATAC	ACTGGTTGAG	6120
	GCCATACTTA	GTAAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTTATGCCTA	TGTTAATGCA	6180
	CTCCTCATTC	CTGGACCAGC	AGGCAAAAAC	AAGCTAGAGA	AACAATTCCA	GCTCCTGAGC	6240
60	CAGTCAAATA	TACAGCAGAG	TGACTATTCT	GCAGCCCTAA	AGCAATGCAA	CAGGGAAAAG	6300
	AATCGAACTT	CTTCTATCAT	CCCTGTGGAA	AGATCAAGGG	TTGGCATTTC	ATCCCTGAGT	6360
	GGAGAAGGCA	CAGACTACAT	CAATGCCTCC	TATATCATGG	GCTATTACCA	GAGCAATGAA	6420
	TTTATCATTA	CCCAGCACCC	TCTCCTTCAT	ACCATCAAGG	ATTTCTGGAG	GATGATATGG	6480
	GACCATAATG	CCCAACTGGT	GGTTATGATT	CCTGATGGCC	AAAACATGGC	AGAAGATGAA	6540
65	TTTGTTTACT	GGCCAAATAA	AGATGAGCCT	ATAAATTTGT	AGAGCTTTAA	GGTCACTCTT	6600
	ATGGCTGAAG	AACACAAATG	TCTATCTAAT	GAGGAAAAAC	TTATAATTCA	GGACTTTATC	6660
	TTAGAAGCTA	CACAGGATGA	TTATGTACTT	GAAGTGAGGC	ACTTTCAGTG	TCCTAAATGG	6720
	CCAAATCCAG	ATAGCCCCAT	TAGTAAACT	TTTGAATTA	TAAGTGTTAT	AAAAGAAGAA	6780
	GCTGCCAATA	GGGATGGGCC	TATGATTGTT	CATGATGAGC	ATGGAGGAGT	GACGGCAGGA	6840
70	ACTTTCCTGT	CTCTGACAAC	CCTTATGCAC	CAACTAGAAA	AAGAAAATTC	CGTGGATGTT	6900
	TACCAGGTAG	CCAAGATGAT	CAATCTGATG	AGGCCAGGAG	TCTTTGCTGA	CATTGAGCAG	6960
	TATCAGTTTC	TCTACAAAGT	GATCCTCAGC	CTTGTGAGCA	CAAGGCAGGA	AGAGAATCCA	7020
	TCCACCTCTC	TGGACAGTAA	TGGTGCAGCA	TTGCCTGATG	GAAATATAGC	TGAGAGCTTA	7080
	GAGTCTTTAG	TTTAACACAG	AAAGGGGTGG	GGGGACTCAC	ATCTGAGCAT	TGTTTTCTCT	7140
75	TTCTTAAAT	TAGGCAGGAA	AATCAGTCTA	GTTCTGTTAT	CTGTTGATTT	CCCATCACCT	7200
	GACAGTAACT	TTTATGACAT	AGGATTCTGC	CGCCAAATTT	ATATCATTTA	CAATGTGTGC	7260
	CTTTTTGCAA	GACTTGTAAT	TTACTTATTA	TGTTTGAAT	AAAATGATTG	AATTTTACAG	7320
	TATTTCTAAG	AATGGAATTG	TGGTATTTTT	TTCTGTATTG	ATTTTAACAG	AAAATTTCAA	7380
	TTTATAGAGG	TTAGGAATTC	CAAACTACAG	AAATTTTGTG	TTTTTAGTGT	CAAATTTTAA	7440
80	GCTGTATTTG	TAGCAATTAT	CAGGTTTGCT	AGAAATATAA	CTTTTAATAC	AGTAGCCTGT	7500
	AAATAAAACA	CTCTTCCATA	TGATATTCAA	CATTTTACAA	CTGCAGTATT	CACCTAAAGT	7560
	AGAAATAATC	TGTTACTTAT	TGTAAATACT	GCCCTAGTGT	CTCCATGGAC	CAAATTTATA	7620
	TTTATAATTG	TAGATTTTAA	TATTTTACTA	CTGAGTCAAG	TTTTCTAGTT	CTGTGTAATT	7680
	GTTTAGTTTA	ATGACGTAGT	TCTTAGCTG	GTCTTACTCT	ACCAAGTTTC	TGACATTGTA	7740
85	TTGTGTTACC	TAAGTCATTA	ACTTTGTTTC	AGCATGTAAAT	TTTAACTTTT	GTGGAAAATA	7800
	GAAATACCTT	CATTTTGAAA	GAAGTTTTTA	TGAGAATAAC	ACCTTACCAA	ACATTGTTCA	7860
	AATGGTTTTT	ATCCAAGGAA	TTGCAAAAAT	AAATATAAAT	ATTGCCATTA	AAAAAAAATA	7920
	AAAAAAAATA	AAAAAAAATA	AAAA				



Seq ID NO: 573 Protein sequence:  
Protein Accession #: Eos sequence

5	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLKKL	KFQGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
	FKASKITFWH	GKCNMSSDGS	EHSLEGQKFP	LEMQIYCFDA	DRFSSFEEAV	KGKGLRLALS	180
10	ILFEVGTEEN	LDFKAIIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVFCEVLTMQ	QSGYVMLMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGKEEIHAEV	CSSEPENVQA	DPENYTSLLV	TWERPRVVDY	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNNLLP	NMSYVLQIVA	ICTNGLYGKY	SDQLIVDMPT	DNPELDLFPE	420
	LIGTEEIIKE	EEEGKDIEEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTTHYNRI	GTKYNEAKTN	480
15	RSPTRGSEFS	GKGDVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPTH	VEGTSASLND	540
	GSKTVLRSPH	MNLSGTAESL	NTVSITEYEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETITYD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NVWFPSSTDI	660
	TAQPDVGSGR	ESFLQNTYTE	IRVDESEKTT	KSFSAAGPMS	QGPSVTDLEM	PHYSTFAYFP	720
	TEVTPHAFTP	SSRQQDLVST	VNVVYSQTTQ	PVYNGETPLQ	PSYSSEVFPL	VTPLLLDNQI	780
20	LNTTPAASSS	DSALHATPVF	PSVDVSFESI	LSSYDGAPLL	PFSSASFSSE	LFRHLHTVSQ	840
	ILPQVTSATE	SDKVPLHASL	PVAGGDLLEL	PSLAQYSDVL	STTHAASETL	EFGESEGVLY	900
	KTLMFQSEVP	PSSDAMMHAR	SSGPEPSYAL	SDNEGSQHIF	TVSYSSAIPV	HDSVGVITYQG	960
	SLFSGPSHIP	IPKSSLITPT	ASLLQPTHAL	SGDGEWSGAS	SDSEFLLPDT	DGLTALNISS	1020
	PVSVAEFTYT	TSVFGDDNKA	LSKSEIIYGN	ETELQIPSFN	EMVYPSESTV	MPNMYDNVVK	1080
25	LNASLQETSV	SISSTKGMFP	GSLAHTTTKV	FDHEISQVPE	NNFSVQPTHT	VSQASGDTSL	1140
	KPVLSANSEP	ASSDPASSEM	LSPSTQLLFY	ETSASFSTEV	LLQPSFQASD	VDTLKTVLP	1200
	AVPSDFILVE	TPKVDKISST	MLHLIVSNSA	SSENMLHSTS	VPVFDVSPTS	HMHSASLQGL	1260
	TISYASEKYE	PVLLKSESSH	QVVPSTLYND	ELFQTANLEI	NQAHPPKGRH	VFATPVLSID	1320
	EPLNTLINKL	IHSDEILTST	KSSVTGKVFA	GIPTVASDTF	VSTDHSVPIG	NGHVAITAVS	1380
30	PHRDGVTST	KLLFPSKATS	ELSHSAKSDA	GLVGGGEDGD	TDDDGDGDDD	DRGSDGLSIH	1440
	KCMSCSSYRE	SQEKVMNDSD	THENSLMDQN	NPISYSLSEN	SEEDNRVTSV	SSDSQTGMDR	1500
	SPGKSPSANG	LSQKHNDGKE	ENDIQTGSAL	LPLSPESKAW	AVLTSDEESG	SGQGTSDSLN	1560
	ENETSTDFSF	ADTNEKDADG	ILAAGDSEIT	PGFFQSPTSS	VTSENSEVFH	VSEAEASNSS	1620
	HESRIGLAEG	LESEKKAVIP	LVIVSALTFI	CLVVLVGILI	YWRKCFQTAH	FYLEDDSTSPR	1680
35	VISTPPTPIF	PISDDVGAIP	IKHFPKHVAD	LHASSGFTEE	FETLKEFYQE	VQSCFVLDGI	1740
	TADSSNHPDN	KHKNRYINIV	AYDHSRVKLA	QLAEKDGLKT	DYINANYVDG	YNRPKAYIAA	1800
	QGPKLSTAED	FWRMIWEHNV	EVIVMITNLV	EKGRRKCDQY	WPADGSEEGY	NFLVTQKSVQ	1860
	VLAYYTVRNF	TLRNTKIKKG	SQKGRPSGRV	VTQYHYTQWP	DMGVPEYSLP	VLTFRVKAAY	1920
	AKRHAVGPVV	VHCSAGVGRT	GTIYIVLDSML	QQIQHEGTVN	IFGFLKHRS	QRNYLVQTEE	1980
40	QYVFHIDTLV	EDSHIKETEV	ALLIPGPAGK	TKLEKQFQLL	SQSNIIQQSDY		2040
	SAALKQCNRE	KNRTSSIIIPV	ERSRVGISSL	SGEGTDYINA	SYIMGYQSN	EFIITQHPLL	2100
	HTIKDFWRMI	WDHNAQLVVM	IPDGQNAED	EFVYWPKNDE	PINCESFKVT	LMAEEHKCLS	2160
	NEEKLIQDF	ILEATQDDYV	LEVRHFQCPK	WPNPDSPISK	TFELISVIKE	EANRNDGPMI	2220
45	VHDEHGGVTA	GTFCALTTLM	HQLEKENSVD	VYQVAKMINL	MRPGVFADIE	QYQFLYKVIL	2280
	SLVSTRQEEN	PSTSLDSNGA	ALPDGNIAES	LESLV			

Seq ID NO: 574 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 148-4518

50	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCACTTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCTCTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
55	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CAGTGAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACCTTAA	TTTCAGGCTT	GGGATAAAAC	ATCATTGGAA	420
60	AACACATTCA	TTCATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTCAGCGGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTTG	720
65	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGAGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAACTCAAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTCTTGTG	AAGTTCTTAC	AATGCAACAA	960
70	TCTGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTCATGA	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACAGCCTT	TCTTGTACAA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATGAGAA	AGTTTGACAT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
75	GGTGCTATTC	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAAAATCA	GGAAAAAGGA	ACCCAGAGTT	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
80	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCAAGTCA	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCTCT	GACTTCTCAG	ACTGTGACTG	AACTGCCACC	TCACACTGTG	1740
	GAAAGTACTT	CAGCTCTTTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCACATATG	1800
	AACTTGTCGG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGTGAAT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
85	GCAACTTCTG	CTATCCCATT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTCTCTCC	1980
	GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAAGAA	TCCTAAAGG	ATCCTTCTAT	GGAGGGAAT	2100



	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCC	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTTGATG	AATCTGAGAA	GACAACCAAG	2220
	TCCTTTTCTG	CAGGCCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
5	CATTATTCTA	CCTTTGCCTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCAACCG	2400
	GTATACAATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAAGGC	AGTTATACCC	CTTGTGATCG	TGTCAGCCCT	GACTTTTATC	2520
	TGTCTAGTGG	TTCTTGTGGG	TATTCCTCAT	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
10	TTTTACTTAG	AGGACAGTAC	ATCCCCCTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTT	2640
	CCAATTTTCA	ATGATGTCGG	AGCAATTCCA	ATAAAGCACT	TTCCAAAGCA	TGTTGCAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTGAAGAA	TTTGAGACAC	TGAAAGAGTT	TTACCAGGAA	2760
	GTGCAGAGCT	GTAAGTTTGA	CTTAGGTATT	ACAGCAGACA	GCTCCAACCA	CCCAGACAAC	2820
	AAGCACAAGA	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAAGCTAGCA	2880
15	CAGCTTGCTG	AAAAGGATGG	CAAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
	TACAACAGAC	CAAAAGCTTA	TATTGCTGCC	CAAGGCCAC	TGAAATCCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAAGTTATT	TCATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAAATG	TGATCAGTAC	TGGCCTGCCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACCTTCTGG	TCACTCAGAA	GAGTGTGCAA	GTGCTTGCC	ATTATACTGT	GAGGAATTTT	3180
20	ACTCTAAGAA	ACACAAAAT	AAAAAAGGGC	TCCCAGAAAG	GAAGACCCAG	TGGACGTGTG	3240
	GTCACACAGT	ATCACTACAC	GCAGTGGCCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGTGAGAAA	GGCAGCCTAT	GCCAAGCGCC	ATGCAGTGGG	GCCTGTGTGC	3360
	GTCCACTGCA	TTGCTGGAGT	TGGAAGAACA	GGCACATATA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTC	AACACGAAGG	AACGTGCAAC	ATATTTGGCT	TCTTAAACA	CATCCGTTC	3480
25	CAAAGAAATT	ATTTGGTACA	AACGTAGGAG	CAATATGTCT	TCATTTCATG	TACACTGGTT	3540
	GAGGCCATAC	TTAGTAAAGA	AACGTAGGTG	CTGCACAGTC	ATATTCATGC	CTATGTAAAT	3600
	GCACTCCTCA	TTCTGGGACC	AGCAGGCAAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGACTAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAATCGAA	CTTCTTCTAT	CATCCCTGTG	GAAAGATCAA	GGGTTGGCAT	TTCATCCCTG	3780
30	AGTGGAGAAG	GCACAGACTA	CATCAATGCC	TCCTATATCA	TGGGCTATTA	CCAGAGCAAT	3840
	GAATTCATCA	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTTCTG	GAGGATGATA	3900
	TGGGACCATA	ATGCCCAACT	GGTGGTTATG	ATTCTGTATG	GCCAAAACAT	GGCAGAAGAT	3960
	GAATTTGTTT	ACTGGCCAAA	TAAAGATGAG	CCTATAAATT	GTGAGAGCTT	TAAGGTCACT	4020
	CTTATGGCTG	AAGAACACAA	ATGCTATCT	AATGAGGAAA	AACTTATAAT	TCAGGACTTT	4080
35	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTGGAAGTGA	GGCACTTTCA	GTGTCTTAAA	4140
	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAAGAA	4200
	GAAGCTGCCA	ATAGGGATGG	GCCTATGATT	GTTTCATGAT	AGCATGGAGG	AGTGACGGCA	4260
	GGAACTTTCT	GTGCTCTGAC	AACCCTTATG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCAGG	TAGCCAAGAT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATTGAG	4380
40	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTTG	GCACAAGGCA	GGAAGAGAAT	4440
	CCATCCACCT	CTCTGGACAG	TAATGGTGCA	GCAATGCTCG	ATGGAAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAAC	CAGAAAGGGG	TGGGGGGACT	CACATCTGAG	CATTGTTTTT	4560
	CTCTTCCTAA	AATTAGGCAG	GAAAATCAGT	CTAGTTCTGT	TATCTGTTGA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTTTCATG	CATAGGATTC	TGCCGCCAAA	TTTATATCAT	TAACATGTG	4680
45	TGCCCTTTTG	CAAGACTTGT	AATTTACTTA	TTATGTTTGA	ACTAAAATGA	TTGAATTTTA	4740
	CAGTATTTCT	AAGAATGGAA	TTGTGGTATT	TTTTTCTGTA	TTGATTTTAA	CAGAAAATTT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAACTA	CAGAAAATGT	TTGTTTTTAG	TGTCAAATTT	4860
	TTAGCTGTAT	TTGTAGCAAT	TATCAGGTTT	GCTAGAAATA	TAACTTTTAA	TACAGTAGCC	4920
	TGTAATAAAA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCAGT	ATTACCTTAA	4980
50	AGTAGAAATA	ATCTGTTACT	TATTGTAAAT	ACTGCCCTAG	TGTCTCCATG	GACCAAATTT	5040
	ATATTTATAA	TTGTAGATTT	TTATATTTTA	CTACTGAGTC	AAGTTTTCTA	GTTCTGTGTA	5100
	ATTGTTTAGT	TTAATGACGT	AGTTCATTAG	CTGGTCTTAC	TCTACCAGTT	TTCTGACATT	5160
	GTATTGTGTT	ACCTAAGTCA	TTAATTTTGT	TTAGCAGATG	AATTTTAACT	TTTGTGGAAA	5220
	ATAGAAATAC	CTTCATTTTG	AAAGAAGTTT	TTAGGAGAAT	AACACCTTAC	CAAACATTGT	5280
55	TCAAATGGTT	TTTATCCAAG	GAATTGCAAA	AATAAATATA	AATATTGCCA	TTAAAAAATA	5340
	AAAAAATAAA	AAAAAATAAA	AAAAAATA				

Seq ID NO: 575 Protein sequence:  
Protein Accession #: Eos sequence

60	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLLKL	KFGQWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
65	FKASKITFWH	GKCNMSDDGS	EHSLEGQKFP	LEMQIYCFDA	DRFSSFEEAV	KGKGLRLALS	180
	ILFEVGTENN	LDFAIIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVFCEVLTMQ	QSGYVLMMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGKEEIHFAV	CSSEPEENVQA	DPENYTSLLV	TWERPRVVYD	TMIEKFAVLY	QQLDGEDQTK	360
70	HEFLTDGYQD	LGAILNNLLP	NMSYVLQIVA	ICTNGLYGKY	SDQLIVDMPT	DNPELDLFPE	420
	LIGTEEIIKE	EEEGKDIEEG	AIVNPGRDSA	TNQRKKEPQ	ISTTHYNNRI	GTKYNEAKTN	480
	RSPTRGSEFS	GKGDVPNTSL	NSTSPVPTKL	ATEKDISLTS	QTVTELPPT	VEGTSASLND	540
	GSKTVLRSPH	MNLSGTAESL	NTVSITEYEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETITYD	VLIPESARNA	SEDSTSSGSE	ESLKDPMEG	NVWFPSSTDI	660
75	TAQPDVSGSR	ESFLQNTYTE	IRVDESEKTT	KSFSAQPVMS	QGPSVTDLEM	PHYSTFAYFP	720
	TEVTPHAFPT	SSRQDDLVT	VNVVYSQTTQ	PVYNAEASNS	SHESRIGLAE	GLESEKKAVI	780
	PLVIVSALTF	ICLVVLVGIL	IYWRKCFQTA	HFYLEDSTSP	RVISTPPTPI	FPISTDDVGA	840
	PIKHFPKHVA	DLHASSGFTE	EPETLKEFYQ	EVQSTVDLG	ITADSSNHPD	NKHKNRYINI	900
	VAYDHSRVKL	AQLAEKDGKL	TDYINANYVD	GYNRPKAYIA	AQGPKLSTAE	DFWRMIWEHN	960
80	VEVIVMITNL	VEKGRRKCDQ	YWPADGSEEE	GNFLVTQKSV	QVLAYYTVRN	FTLRNTKIKK	1020
	GSQKGRPSGR	VVTQYHYTQW	PDMGVPEYSL	PVLTFVRKAA	YAKRHAVGPV	VVHCSAGVGR	1080
	TGTYIVLDSM	LQQIQHEGTV	NIFGFLKHIR	SQRNYLVQTE	EQYVFIHDTL	VEAILSKETE	1140
	VLDSHIHAYV	NALLIPGPAG	KTKLEKQFQL	LSQSNIIQSD	YSAALKQCNR	EKNRTSSIIP	1200
	VERSRVGISS	LSGEGTDYIN	ASYIMGYQVS	NEFIITQHPL	LHTIKDFWRM	IWDHNAQLVV	1260
85	MIPDQGNMAE	DEFVYWPND	EPINCESYKV	FLMAEEHKCL	SNEEKLIIQD	FILEATQDDY	1320
	VLEVRHFQCP	KWPNPDSPIS	KTFELISVIK	EEAANRDGPM	IVHDEHGGVT	AGTFCALTTL	1380
	MHOLEKENS	DVYQVAKMIN	LMRPGVFADI	EQYQFLYKVI	LSLVSTRQEE	NPSTSLDSNG	1440
	AALPDGNIAE	SLESILV					



Seq ID NO: 576 DNA sequence  
Nucleic Acid Accession #: EOS sequence  
Coding sequence: 148-4494

5	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCATTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAG	ATTTCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
10	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTTGAA	420
15	AACACATTCA	TTCATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTCAGCGGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCAGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
20	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTAGTTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAACTCTCA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGCATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTTTGTG	AAGTTCTTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAAG	1020
25	TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCATGA	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAG	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTC	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
30	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAAAATCA	GGAAAAGGGA	ACCCAGATTG	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	TCCCAACCGA	TCCCAACCAA	GAGGAAGTGA	ATTCTCTGGA	1620
35	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCAGTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCTCT	GACTTCTCAG	ACTGTGACTG	AACTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAAGTG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTCCG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
40	GCAACTCTTG	CTATCCCATT	CATCTGTAGT	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTTCAAGAGG	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
45	TCCTTTTCTG	CAGGCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CCTTTGCCTA	CTTCCCAACT	GAGGTAAACG	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTGGT	CTCCACGGTC	AACCTGGTAT	ACTCGCAGAC	AACCAACCCG	2400
	GTATACAATG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
50	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
	ATTTAGATG	ATGTCGGAGC	AATCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGGAAGTGC	AGAGCTGTAC	TGTTGACTTA	2760
	GGTATTACAG	CAGACAGCTC	CAACCACCCA	GACAACAAGC	ACAAGAATCG	ATACATAAAT	2820
55	ATCGTTGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCAAA	2880
	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
	GCTGCCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTTCT	GGAGAATGAT	ATGGGAACAT	3000
	AATGTGGAAG	TTATTGTCAT	GATAACAAC	CTCGTGAGAG	AAGGAAGGAG	AAAATGTGAT	3060
	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAC	TTCTGGTCAC	TCAGAAGAGT	3120
60	GTGCAAGTGC	TTGCCTATTA	TACTGTGAGG	AATTTTACTC	TAAGAAACAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGA	CGTGTGGTCA	CACAGTATCA	CTACACGCAG	3240
	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGGCA	3300
	GCCTATGCCA	AGCGCCATGC	AGTGGGGCCT	GTTGTGCTCC	ACTGCAGTGC	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAACT	3420
65	GTCAACATAT	TTGGCTTCTT	AAAACACATC	CGTTACACAA	GAAATTATTT	GGTACAAACT	3480
	GAGGAGCAAT	ATGTCTTCAT	TCATGATACA	CTGGTTGAGG	CCATACTTAG	TAAAGAAACT	3540
	GAGGTGCTGG	ACAGTCATAT	TCATGCCTAT	GTTAATGCAC	TCCTCATTCC	TGGACCAGCA	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCCAG	CTCCTGAGCC	AGTCAAATAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAAGA	ATCGAACTTC	TTCTATCATC	3720
70	CCTGTGGAAG	GATCAAGGGT	TGGCATTTC	TCCCTGAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGCCCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCATTAC	CCAGCACCCCT	3840
	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	ACCATAATGC	CCAACGGGTG	3900
	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACTCTTA	TGGCTGAAGA	ACACAAATGT	4020
75	CTATCTAATG	AGGAAAAACT	TATAATTTCAG	GACTTTATCT	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	CAAATCCAGA	TAGCCCCATT	4140
	AGTAAACTTT	TTGAACCTAT	AAGTGTTATA	AAAGAAGAAG	CTGCCAATAG	GGATGGGCCT	4200
	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	CTTCTGTGTC	TCTGACAACT	4260
	CTTATGCACC	AACTAGAAAA	AGAAAATTCC	GTGGATGTTT	ACCAGGTAGC	CAAGATGATC	4320
80	AATCTGATGA	GGCCAGGAGT	CTTTGCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACCTCTCT	GGACAGTAAT	4440
	GGTGCAGCAT	TGCCTGATGG	AAATATAGCT	GAGAGCTTAG	AGTCTTTAGT	TTAACACAGA	4500
	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTTCCCTCT	TCCTAAAATT	AGGCAGGAAA	4560
	ATCAGTCTAG	TTTGTATTAT	TGTTGATTTC	CCATCACCTG	ACAGTAACTT	TCATGACATA	4620
85	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	TTTTTGCAAG	ACTTGTAAAT	4680
	TACTTATTAT	GTTTGAACCT	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
	GGTATTTTTT	TCTGTATTGA	TTTTAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTCC	4800



AAACTACAGA AAATGTTTGT TTTTAGTGTC AAATTTTTAG CTGTATTTGT AGCAATTATC 4860  
 AGGTTTGCTA GAAATATAAC TTTTAATACA GTAGCCTGTA AATAAACAC TCTTCCATAT 4920  
 GATATTCAAC ATTTTACAAC TGCAGTATTC ACCTAAAGTA GAAATAATCT GTTACTTATT 4980  
 GTAAATACTG CCCTAGTGTC TCCATGGACC AAATTTATAT TTATAATTGT AGATTTTTAT 5040  
 ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTGTAATTG TTTAGTTTAA TGACGTAGTT 5100  
 CATTAGCTGG TCTTACTCTA CCAGTTTTCT GACATTGTAT TGTGTTACCT AAGTCATTAA 5160  
 CTTTGTTC A GCATGTAATT TTAACCTTTG TGGAAAATAG AAATACCTTC ATTTTGAAAG 5220  
 AAGTTTTTAT GAGAATAACA CCTTACCAA CATTTGTTCAA ATGGTTTTTA TCCAAGGAAT 5280  
 TGCAAAAATA AATATAAATA TTGCCATTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 5340  
 AAA

Seq ID NO: 577 Protein sequence:  
 Protein Accession #: EOS sequence

1 11 21 31 41 51  
 | | | | | |  
 MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKKL KFQGWDKTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFWH GKCNMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEEAV KGKGLRLALS 180  
 ILFEVGTEEN LDFKAIIDGV ESVSFRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVMLMDY LQNNFREQQY KFSRQVFSSY 300  
 TGKEEIEHAV CSSEPEENVQA DPENYTSLLV TWERPRVVD TMIEKFAVLY QQLDGEDQTK 360  
 HEFLTDGYQD LGAILNLLP NMSYVLQIVA ICTNGLYGKY SDQLIVDMPT DNPEDLDFPE 420  
 LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDPVNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPHT VEGTSASLND 540  
 GSKTVLRSPH MNLSTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSMEG NVWFPSSTDI 660  
 TAQPDVSGSR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTPHAFTP SSRQDLVST VNVVYSQTTQ PVYNEASNSS HESRIGLAEG LESEKKAVIP 780  
 LVIVSALTFI CLVVVLGILI YWRKCFQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840  
 IKHFPKHVAD LHASSGFTEE FEEVQSCTVD LGITADSSNH PDKHKNRYI NIVAYDHSRV 900  
 KLAQLAEKDG KLTDYINANY VDGYNRPKAY IAAQGPLKST AEDFWRMIWE HNVEIVMIT 960  
 NLVEKGRKRC DQYWPADGSE EYGNFLVTQK SVQVLAYYTV RNFTLRNTKI KKGSKGRPS 1020  
 GRVVTQYHYT QWPDMGVPEY SLPVLTFRK AAYAKRHAVG PVVHCSAGV GRTGTIYVLD 1080  
 SMLQQIQHEG TVNIFGFLKH IRSQRNYLVQ TEEQYVFIHD TLVEAILSKE TEVLDSHIHA 1140  
 YVNALLIPGP AGKTKLEKQF QLLSQSNIQQ SDYSAALKQC NREKNRTSSI IPVERSRVGI 1200  
 SSLSGEGTDY INASYIMGYQ QSNEFIITQH PLLHTIKDFW RMIWDHNAQL VVMIPDQNM 1260  
 AEDEFVYWPN KDEPINCESF KVTLMAEHK CLSNEEKLI QDFILEATQD DYVLEVRHFQ 1320  
 CPKWPNDPSP ISKTFELISV IKEEAANRDG PMIVHDEHGG VTAGTFCALT TLMHQLEKEN 1380  
 SVDVYQVAKM INLMRPGVFA DIEQYQFLYK VILSLVSTRQ BENPSTSLDS NGAALPDGNI 1440  
 AESLESLV

Seq ID NO: 578 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 501-4514

1 11 21 31 41 51  
 | | | | | |  
 CACACATACG CACGCACGAT CTCACTTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAAC ATTTCTCTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGCGAGGGG CCGCAGACCG TCTGGAATG GCTATCTTAA AGCGTTTCCG CGTTTGCAAT 180  
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTTGTTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAAT TGGGGAAAGA 300  
 AATATCCAAC ATGTAATAGC CCAAAACAAT CTCCTATCAA TATTGATGAA GATCTTACAC 360  
 AAGTAAATGT GAATCTTAAG AACTTTAAAT TTCAGGGTTG GGATAAAACA TCATTGGAAA 420  
 ACACATTCAT TCATAACACT GGGAAAACAG TGGAAATTA TCTCACTAAT GACTACCGTG 480  
 TCAGCGGAGG AGTTTCAGAA ATGGTGTGTT AAGCAAGCAA GATAACTTTT CACTGGGGAA 540  
 AATGCAATAT GTCATCTGAT GGATCAGAGC ATAGTTTAGA AGGACAAAAA TTTCCACTTG 600  
 AGATGCAAAAT CTACTGCTTT GATGCGGACC GATTTTCAAG TTTTGAGGAA GCAGTCAAAG 660  
 GAAAAGGGAA GTTAAGAGCT TTATCCATTT TGTGTTGAGG TGGGACAGAA GAAAATTTGG 720  
 ATTTCAAAGC GATTATTGAT GGAGTCGAAA GTGTTAGTCG TTTTGGAAG CAGGCTGCTT 780  
 TAGATCCATT CATACTGTTG AACCTTCTGC CAACTCAAC TGACAAGTAT TACATTTACA 840  
 ATGGCTCATT GACATCTCCT CCCTGCACAG ACACAGTTGA CTGGATTGTT TTAAAGATA 900  
 CAGTTAGCAT CTCTGAAAGC CAGTTGGCTG TTTTGTGTA AGTTCTTACA ATGCAACAAT 960  
 CTGGTTATGT CATGCTGATG GACTACTTAC AAAACAATTT TCGAGAGCAA CAGTACAAGT 1020  
 TCTCTAGACA GGTGTTTTCC TCATACACTG GAAAGGAAGA GATTCATGAA GCAGTTTGTA 1080  
 GTTCAAGACC AGAAAATGTT CAGGCTGACC CAGAGAATTA TACCAGCCTT CTGTTACAT 1140  
 GGGAAAGACC TCGAGTCGTT TATGATACCA TGATTGAGAA GTTTGCAGTT TTGTACCAGC 1200  
 AGTTGGATGG AGAGGACCAA ACCAAGCATG AATTTTGTAC AGATGGCTAT CAAGACTTGG 1260  
 GTGCTATTCT CAATAATTG CTACCCAATA TGAGTTATGT TCTTCAGATA GTAGCCATAT 1320  
 GCACTAATGG CTTATATGGA AAATACAGCG ACCAAGTATG TGTCGACATG CCTACTGATA 1380  
 ATCTGAACT TGATCTTTTC CCTGAATTAA TTGGAAGTGA AGAAATAATC AAGGAGGAGG 1440  
 AAGAGGGAAA AGACATTGAA GAAGGCGCTA TTGTGAATCC TGGTAGAGAC AGTGTACAA 1500  
 ACCAAATCAG GAAAAGGAA CCCCAGATT CTACCACAAC AACTACAAT CGCATAGGGA 1560  
 CGAAATACAA TGAAGCCAAG ACTAACCAGT CCCCACAAG AGGAAGTGAA TTCTCTGGAA 1620  
 AGGGTGATGT TCCCAATACA TCTTTAAATT CCAGTTCCCA ACCAGTCACT AAATTAGCCA 1680  
 CAGAAAAAGA TATTTCTTGG ACTTCTCAGA CTGTGACTGA ACTGCCACCT CACTGTGG 1740  
 AAGGTACTTC AGCCTCTTTA AATGATGGCT CTAAGACTGT TCTTAGATCT CCACATATGA 1800  
 ACTTGTGGG GACTGCAGAA TCCTTAAATA CAGTTTCTAT AACAGAATAT GAGGAGGAGA 1860  
 GTTTATTGAC CAGTTTCAAG CTTGATACTG GAGCTGAAGA TTCTTCAGGC TCCAGTCCCG 1920  
 CAACTTCTGC TATCCCATTC ATCTCTGAGA ACATATCCCA AGGGTATATA TTTTCTCCG 1980  
 AAAACCCAGA GACAATAACA TATGATGTC TTATACCAGA ATCTGCTAGA AATGCTTCCG 2040  
 AAGATTCAAC TTCATCAGGT TCAGAAAGAT CACTAAAGGA TCCTTCTATG GAGGGAAATG 2100  
 TGTGGTTTCC TAGCTCTACA GACATAACAG CACAGCCCGA TGTTGGATCA GGCAGAGAGA 2160  
 GCTTCTCCA GACTAATTAC ACTGAGATAC GTGTTGATGA ATCTGAGAAG ACAACCAAGT 2220  
 CTTTCTCTGC AGGCCAGTG ATGTCACAGG GTCCCTCAGT TACAGATCTG GAAATGCCAC 2280



ATTATTCTAC CTTTGCCTAC TTCCCAACTG AGGTAACACC TCATGCTTTT ACCCCATCCT 2340  
 CCAGACAACA GGATTTGGTC TCCACGGTCA ACGTGGTATA CTCGCAGACA ACCCAACCGG 2400  
 TATACAATGA GGCCAGTAAT AGTAGCCATG AGTCTCGTAT TGGTCTAGCT GAGGGGTTGG 2460  
 AATCCGAGAA GAAGGCAGTT ATACCCCTTG TGATCGTGTC AGCCCTGACT TTTATCTGTC 2520  
 TAGTGGTTCT TGTGGGTATT CTCATCTACT GGAGGAAATG CTTCCAGACT GCACACTTTT 2580  
 ACTTAGAGGA CAGTACATCC CCTAGAGTTA TATCCACACC TCCAACACCT ATCTTTCCAA 2640  
 TTTCAAGATGA TGTCGGAGCA ATTCCAATAA AGCACTTTCC AAAGCATGTT GCAGATTTAC 2700  
 ATGCAAGTAG TGGGTTTACT GAAGAATTTG AGACACTGAA AGAGTTTAC CAGGAAGTGC 2760  
 AGAGCTGTAC TGTTGACTTA GGTATTACAG CAGACAGCTC CAACCACCCA GACAACAAGC 2820  
 ACAAGAATCG ATACATAAAT ATCGTTGCCT ATGATCATAG CAGGGTTAAG CTAGCACAGC 2880  
 TTGCTGAAAA GGATGGCAAA CTGACTGATT ATATCAATGC CAATTATGTT GATGGCTACA 2940  
 ACAGACCAAA AGCTTATATT GCTGCCCCAAG GCCCACTGAA ATCCACAGCT GAAGATTTCT 3000  
 GGAGAATGAT ATGGGAACAT AATGTGGAAG TTATTGTTCAT GATAACAAAC CTCGTGGAGA 3060  
 AAGGAAGGAG AAAATGTGAT CAGTACTGGC CTGCCGATGG GAGTGAGGAG TACGGGAAC 3120  
 TTCTGGTCAC TGCAAGAGT GTGCAAGTGC TTGCTATTA TACTGTGAGG AATTTTACTC 3180  
 TAAGAAACAC AAAAATAAAA AAGGGCTCCC AGAAAGGAAG ACCCAGTGGA CGTGTGGTCA 3240  
 CACAGTATCA CTACACGCAG TGGCCTGACA TGGGAGTACC AGAGTACTCC CTGCCAGTGC 3300  
 TGACCTTTGT GAGAAAGGCA GCCTATGCCA AGCGCCATGC AGTGGGGCCT GTTGTGCTCC 3360  
 ACTGCAGTGC TGGAGTTGGA AGAACAGGCA CATATATTGT GCTAGACAGT ATGTTGCAGC 3420  
 AGATTCAACA CGAAGGAACT GTCAACATAT TTGGCTTCTT AAAACACATC CGTTCACAAA 3480  
 GAAATTATTT GGTAGCAACT GAGGAGCAAT ATGTCTTCAT TCATGATACA CTGGTTGAGG 3540  
 CCATACTTAG TAAAGAAACT GAGGTGCTGG ACAGTCATAT TCATGCCTAT GTTAATGCAC 3600  
 TCCTCATTCC TGGACCAGCA GGCAAAACAA AGCTAGAGAA ACAATTCCAG CTCCTGAGCC 3660  
 AGTCAAATAT ACAGCAGAGT GACTATTCTG CAGCCCTAAA GCAATGCAAC AGGGAAAAGA 3720  
 ATCGAACTTC TTCTATCATC CCTGTGGAAG GATCAAGGGT TGGCATTTC TCCCTGAGTG 3780  
 GAGAAGGCAC AGACTACATC AATGCCTCCT ATATCATGGG CTATTACCAG AGCAATGAAT 3840  
 TCATCATTAC CCAGCACCCCT CTCCTTTCAT CCATCAAGGA TTTCTGGAGG ATGATATGGG 3900  
 ACCATAATGC CCAACTGGTG GTTATGATT C TGATGGCCA AAACATGGCA GAAGATGAAT 3960  
 TTGTTTACTG GCCAAATAAA GATGAGCCTA TAAATTGTGA GAGCTTTAAG GTCACCTCTA 4020  
 TGGCTGAAGA ACACAAATGT CTATCTAATG AGGAAAACT TATAATTCAG GACTTTATCT 4080  
 TAGAAGCTAC ACAGGATGAT TATGTACTTG AAGTGAGGCA CTTTCAGTGT CCTAAATGGC 4140  
 CAAATCCAGA TAGCCCCATT AGTAAACTTT TTGAACCTAT AAGTGTATA AAAGAAGAAG 4200  
 CTGCCAATAG GGATGGGCCT ATGATTGTTC ATGATGAGCA TGGAGGAGTG ACGGCAGGAA 4260  
 CTTTCTGTGC TCTGACAACC CTTATGCACC AACTAGAAAA AGAAAAATTC GTGGATGTTT 4320  
 ACCAGGTAGC CAAGATGATC AATCTGATGA GGCCAGGAGT CTTTGCTGAC ATTGAGCAGT 4380  
 ATCAGTTTCT CTACAAAGTG ATCCTCAGCC TTGTGAGCAC AAGGCAGGAA GAGAATCCAT 4440  
 CCACCTCTCT GGACAGTAAT GGTGCAGCAT TGCCTGATGG AAATATAGCT GAGAGCTTAG 4500  
 AGTCTTTAGT TTAACACAGA AAGGGGTGGG GGGACTCACA TCTGAGCATT GTTTTCCTCT 4560  
 TCCTAAAAAT AGGCAGGAAA ATCAGTCTAG TTCTGTTATC TGTTGATTTC CCATCACCTG 4620  
 ACAGTAACCT TCATGACATA GCAATCTGCC GCAAAATTTA TATCATTAA AATGTGTGCC 4680  
 TTTTGTCAAG ACTTGTAATT TACTTATTAT GTTGAACCTA AAATGATTGA ATTTTACAGT 4740  
 ATTTCTAAGA ATGGAATTGT GGTATTTTTT TCTGTATTGA TTTTAACAGA AAATTTCAAT 4800  
 TTATAGAGGT TAGGAATTCC AAACACAGA AAATGTTTGT TTTTAGTGTC AAATTTTATG 4860  
 CTGTATTTGT AGCAATTATC AGGTTTGCTA GAAATATAAC TTTTAATACA GTAGCCTGTA 4920  
 AATAAAACAC TCTTCCATAT GATATTCAAC ATTTTACAAC TGCAGTATTC ACCTAAAGTA 4980  
 GAAATAATCT GTTACTTATT GTAAATCATG CCCTAGTGTC TCCATGGACC AAATTTATAT 5040  
 TTATAATTGT AGATTTTAT ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTGTAATTG 5100  
 TTTAGTTTAA TGACGTAGTT CATTAGCTGG TCTTACTCTA CCAGTTTTCT GACATTGTAT 5160  
 TGTGTTACCT AAGTCATTAA CTTTGTTCAT GCATGTAATT TTAACTTTGT TGGAAAATAG 5220  
 AAATACCTTC ATTTTGAAAG AAGTTTTTAT GAGAAATAACA CCTTACCAAA CATTGTTCAA 5280  
 ATGGTTTTTA TCCAAGGAAT TGCAAAAATA AATATAAATA TTGCCATTAA AAAAAAAAAA 5340  
 AAAAAAAAAA AAAAAAAAAA AAA

Seq ID NO: 579 Protein sequence:  
 Protein Accession #: EOS sequence

1 11 21 31 41 51  
 | | | | |  
 MVFKASKITF HWGKCNMSSD GSEHSLEGQK FPLEMQIYCF DADRFSSFEE AVKGKGLRA 60  
 LSILFEVGT E NLDPKAIID GVESVSRFGK QAALDPFILL NLLPNSTDYK YIYNGSLTSP 120  
 PCTDVTVDWIV FKDTVSISES QLAVFCEVLT MQQSGYVMLM DYLNQNFREQ QYKFSRQVFS 180  
 SYTGKEEIEH AVCSSEPEENV QADPENYTS LVTWERPRVV YDTMIEKFAV LYQQLDGEDQ 240  
 TKHEFLTDGY QDLGAILNNL LPNMSYVLQI VAICTNGLYG KYSDQLIVDM PTDNPELDLF 300  
 PELIGTEEII KEEEGKDIE EGAIVNPGRD SATNQIRKKE PQISTTTHYN RIGTKYNEAK 360  
 TNRSPTRGSE FSGKGDVPNT SLNSTSPQVT KLATEKDISE TSQTVTELP HTVEGTSASL 420  
 NDGSKTVLRS PHMNLSTAE SLNTVSITEY EEESLLTSFK LDTGAEDSSG SSPATSAIPF 480  
 ISENISQGYI PSSENPETIT YDVLIPESAR NASEDSTSSG SEESLKDPSM EGNVWFPSST 540  
 DITAQPDVGS GRESFLQTN Y TEIRVDESEK TTKSFSAGPV MSQGPSVTDL EMPHYSTFAY 600  
 FPTEVTPHAF TPSSRQQDLV STVNVSYSQT TQPVYNEASN SSHESRIGLA EGLESEKKAV 660  
 IPLVIVSALT FICLVVLVGI LIYWRKCFQT AHFYLEDSTS PRVISTPPTP IFPISDDVGA 720  
 IPIKHPKHKV ADLHASSGFT EEFETLKEFY QEVQSCTVDL GITADSSNHP DNKHKRYIN 780  
 IVAYDHSRVK LAQLAEKDGK LTDYINANYV DGYNRPKAYI AAQGPLKSTA EDFWRMIWEH 840  
 NVEVIVMITN LVEKGRRKCD QYWPADGSEE YGNFLVTQKS VQVLAYYTVR NFTLRNTKIK 900  
 KGSQKGRPSG RVVTQYHYTQ WPDGMVPEYS LPVLTFRKA AYAKRHAVGP VVVHCSAGVG 960  
 RTGTYIVLDS MLQQIQHEGT VNIFGFLKHI RSQRNYLVQT EEQYVFIHDT LVEAILSKET 1020  
 EVLDSHIHAY VNALLIPGPA GKTKLEKQFQ LLSQSNIQPS DYSALKQCN REKNRTSSII 1080  
 PVERSRVGIS SLSGEGTDYI NASYIMGYQ SNEFIITQHP LLHTIKDFWR MIWDHNAQLV 1140  
 VMIPDGQNMMA EDEFVYWPKNK DEPINCSEFK VTLMAEEHKC LSNEEKLIQ DFILEATQDD 1200  
 YVLEVRHFQC PKWPNPDSPI SKTFELISVI KEEAANRDGP MIVHDEHGGV TAGTFCALTT 1260  
 LMHQLEKENS VDVYQVAKMI NLMRPGVFAD IEQYQFLYKV ILSLVSTRQE ENPSTSLDSN 1320  
 GAALPDGNIA ESLESILV

Seq ID NO: 580 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 148-4632

1 11 21 31 41 51



	CACACATACG	CACGCACGAT	CTCACCTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAG	ATTTCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
5	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AACGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
10	AACACATTCA	TTCATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTCAGCGGAG	GAGTTTCAGA	AATGGTGTGT	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
15	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTTTGTG	AAGTTCTTAC	AATGCAACAA	960
	TCTGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAA	1020
20	TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCATGA	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTC	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
25	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAAATCA	GGAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
30	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCAGTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCTTT	GACTTCTCAG	ACTGTGACTG	AACTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACTG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTCGG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
35	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAAGAA	TCATAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCC	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
40	TCCTTTCTG	CAGGCCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATCTA	CCTTTGACCT	CTTCCCAACT	GAGGTAAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCAACCG	2400
	GTATACAATG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
45	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGTTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTA	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
	ATTTAGATG	ATGTGCGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	2760
	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	GCAGACAGCT	CCAACCACCC	AGACAACAAG	2820
50	CACAAGAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	2880
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	2940
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTT	3000
	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTTAATGTCA	TGATAACAAA	CCTCGTGGAG	3060
	AAAGGAAGGA	GAAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	3120
55	TTTCTGGTCA	CTCAGAAGAG	TGTGCAAGTG	CTTGCCATT	ATACTGTGAG	GAATTTTACT	3180
	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	3240
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	3300
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTCTGC	3360
	CACTGCAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	3420
60	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAAACACAT	CCGTTTACAA	3480
	AGAAATTAT	TGGTACAAAC	TGAGGAGCAA	TATGTCTTCA	TTTATGATAC	ACTGGTTGAG	3540
	GCCATACTTA	GTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTTATGCCTA	TGTTAATGCA	3600
	CTCCTCATTC	CTGGACCAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GGGTCTCACT	3660
	CTGTCAACCA	GGCTGGAGTG	CAGAGGCACA	ATCTCGGCTC	ACTGCAACCT	TCCTCTCCCT	3720
65	GGCTTAACCTG	ATCCTCCTAC	CTCAGCTTCC	CGAGTGGCTG	GGACTATACT	CCTGAGCCAG	3780
	TCAAATATAC	AGCAGAGTGA	CTATTCTGCA	GGCTTAAAGC	AATGCAACAG	GGAAAAGAA	3840
	CGAACTTCTT	CTATCATCCC	TGTGGAAGAA	TCAAGGGTTG	GCATTTTCATC	CCTGAGTGGA	3900
	GAAGGCACAG	ACTACATCAA	TGCCTCCTAT	ATCATGGGCT	ATTACCAGAG	CAATGAATTC	3960
	ATCATTACCC	AGCACCTCT	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	4020
70	CATAATGCCC	AACTGGTGGT	TATGATTCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	4080
	GTTTACTGGC	CAAATAAAGA	TGAGCCTATA	AATTGTGAGA	GCTTTAAGGT	CACTCTTATG	4140
	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTA	TAATTCAGGA	CTTTATCTTA	4200
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTTATGTCTC	TAAATGGCCA	4260
	AATCCAGATA	GCCCCATTAG	TAAAACTTTT	GAATTTATAA	GTGTTATAAA	AGAAGAAGCT	4320
75	GCCAATAGGG	ATGGGCCTAT	GATTGTTTCT	GATGAGCATG	GAGGAGTGAC	GGCAGGAACT	4380
	TTCTGTGCTC	TGACAACCTT	TATGCACCAA	CTAGAAAAAG	AAAATTCCGT	GGATGTTTAC	4440
	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	4500
	CAGTTTCTCT	ACAAAGTGAT	CCTCAGCCTT	GTGGGCACAA	GGCAGGAAGA	GAATCCATCC	4560
80	ACCTCTCTGG	ACAGTAATGG	TGCAGTATTG	CCTGAGTGAA	ATATAGCTGA	GAGCTTAGAG	4620
	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	4680
	CTAAAAATTAG	GCAGGAAAAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCC	ATCACCTGAC	4740
	AGTAACTTTC	ATGACATAGG	ATTCTGCCGC	CAAAATTATA	TCATTAACAA	TGTGTGCCCT	4800
	TTTGCAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACTAAA	ATGATTGAAT	TTTACAGTAT	4860
	TTCTAAGAA	GGAATTGTGG	TATTTTTTTT	TGTATTGATT	TTAACAGAAA	ATTTCAATTT	4920
85	ATAGAGGTTA	GGAATTCCAA	ACTACAGAAA	ATGTTTGTTC	TTAGTGTCAA	ATTTTGTAGCT	4980
	GTATTTGTAG	CAATTATCAG	GTTTGCTAGA	AATATAACTT	TTAATACAGT	AGCCTGTAAA	5040
	TAAACACTC	TTCCATATGA	TATTCACAT	TTTACAACCT	CAGTATTCAC	CTAAAGTAGA	5100
	AATAATCTGT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	5160



5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65  
 70  
 75  
 80  
 85

ATAATTGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTTCTG TGTAAATTGTT 5220  
 TAGTTTAATG ACGTAGTTCA TTAGCTGGTC TTAGCTTACC AGTTTCTGA CATTGTATTG 5280  
 TGTACCTAA GTCATTAACT TTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAAAAGAA 5340  
 ATACCTTCAT TTTGAAAGAA GTTTTATGA GAATAACACC TTACCAACA TTGTTCAAAT 5400  
 GGTTTTATC CAAGGAATTG CAAAAATAAA TATAAATATT GCCATTAAAA AAAAAA 5460  
 AAAAAA AAAAAA A

Seq ID NO: 581 Protein sequence:  
 Protein Accession #: EOS sequence

1 11 21 31 41 51  
 | | | | |  
 MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKKL KFQGWKDTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFWH GKCNMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEEAV KGKGLRLALS 180  
 ILFEVGTEN LDFKAIIDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVLMMDY LQNNFREQQY KFSRQVFSSY 300  
 TGKEEIEHAV CSSEPEENVQA DPENYTSLLV TWERPRVVD TMIEKFAVLY QQLDGEDQTK 360  
 HEFLTDGYQD LGAILNNLLP NMSYVLQIVA ICTNGLYGKY SDQLIVDMPT DNPELDLFPE 420  
 LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPT VEGTSASLND 540  
 GSKTVLRSPH MNLSGTAESL NTVSITEYEE ESSLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSMEG NVWFPSSTDI 660  
 TAQPDVSGSR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTPHAFTP SSRQQDLVST VNVVYSQTTQ PVYNEASNSS HESRIGLAEG LESEKKAVIP 780  
 LVIVSALTPI CLVVLVGILI YWRKCFQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840  
 IKHFPKHVAD LHASSGFTEE FETLKEFYQE VQSCTVDLGI TADSSNHPDN KHKNRYINIV 900  
 AYDHSRVKLA QLAEKDGKLT DYINANYVDG YNRPKAYIAA QGPLKSTAED FWRMIWEHNV 960  
 EVIVMITNLV EKGRRKCDQY WPADGSEYEG NFLVTQKSVQ VLAYYTVRNF TLRNTKIKKG 1020  
 SQKGRPSGRV VTQYHYTQWP DMGVPEYSLP VLTFRKAAAY AKRHAVGPVV VHCSAGVGRT 1080  
 GTYIVLDSML QQIQHEGTVN IFGFLKHIRS QRNYLVQTEE QYVFIHDTLV EAILSKEDEV 1140  
 LDSHIHAYVN ALLIPGPAGK TKLEKQFQGL TSPRLBCRG TISAHCNLP PGLTDPPTSA 1200  
 SRVAGTILLS QSNIQQSDYS AALKQCNRK NRTSSIIPVE RSRVGISSLS GEGTDYINAS 1260  
 YIMGYQSNE FIITQHPLLH TIKDFWRMIW DHNAQLVVM PDGQNMMAEDE FVYWPKNDEP 1320  
 INCESFKVTL MAEEHKCLSN EEKLIQDFI LEATQDDYVL EVRHFQCPKW PNPDSPIST 1380  
 FELISVIKEE AANRDGPMIV HDEHGGVTAG TFCALTLMH QLEKENSVDV YQVAKMINLM 1440  
 RPFVADIEQ YQFLYKVILS LVGTRQEENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 582 DNA sequence  
 Nucleic Acid Accession #: NM\_002851.1  
 Coding sequence: 148..7092

1 11 21 31 41 51  
 | | | | |  
 CACACATACG CACGCACGAT CTCACTTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAGC ATTTCCCTTC CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCCT CGCTTGCAAT 180  
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTTGTTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAA TTGGGGAAAG 300  
 AAATATCCAA CATGTAATAG CCCAAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA 360  
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA 420  
 AACACATCA TTCATAACAC TGGGAAAACA GTGGAAATTA ATCTCACTAA TGACTACCGT 480  
 GTCAGCGGAG GAGTTTCAGA AATGGTGTTC AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 AAATGCAATA TGTCATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600  
 GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAA 660  
 GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAATTTG 720  
 GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTC GTTTTGGGAA GCAGGCTGCT 780  
 TTAGATCCAT TCATACTGTT GAACCTTCTG CCAACTCAA CTGACAAGTA TTACATTTAC 840  
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTTTGTG AAGTTCTTAC AATGCAACAA 960  
 TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTCGAGAGCA ACAGTACAAG 1020  
 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTCATGA AGCAGTTTGT 1080  
 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA 1140  
 TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCAGT TTTGTACCAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 TGCATAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380  
 AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAACTG AAGAAATAAT CAAGGAGGAG 1440  
 GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500  
 AACCAATCA GGAAAAAGGA ACCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560  
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620  
 AAGGGTGATG TTCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTAC TAAATTAGCC 1680  
 ACAGAAAAAG ATATTTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAACTG TTCTTAGATC TCCACATATG 1800  
 AACTTGTCGG GGAAGTCAGA ATCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860  
 AGTTTATTGA CCAAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920  
 GCAACTTCTG CTATCCCATC CATCTCTAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980  
 GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040  
 GAAGATTCAA CTTATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAT 2100  
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTTCTCC AGACTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220  
 TCCTTTTCTG CAGGCCAGT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA 2280  
 CATTATTCTA CCTTGCCTA CTTCCCACT GAGGTAAAC CTCTATGCTTT TACCCATCC 2340  
 TCCAGACAAC AGGATTGGT CTCCACGTC AACGTGTAT ACTCGCAGAC AACCCAACCG 2400  
 GTATACAATG GTGAGACACC TCTTCAACCT TCCTACAGTA GTGAAGTCTT TCCTCTAGTC 2460  
 ACCCCTTTGT TGCTTGACAA TCAGATCCTC AACACTACCC CTGCTGCTTC AAGTAGTGAT 2520



	TCGGCCTTGC	ATGCTACGCC	TGTATTTCCC	AGTGTGCGATG	TGTCATTTGA	ATCCATCCTG	2580
	TCTTCCTATG	ATGGTGCACC	TTTGCTTCCA	TTTTCCTCTG	CTTCCTTCAG	TAGTGAATTG	2640
	TTTCGCCATC	TGCATACAGT	TTCTCAAATC	CTTCCACAAG	TTACTTCAGC	TACCGAGAGT	2700
5	GATAAGGTGC	CCTTGCATGC	TTCTCTGCCA	GTGGCTGGGG	GTGATTTGCT	ATTAGAGCCC	2760
	AGCCTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGGTAGTG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACCACCC	2880
	AGCAGTGATG	CCATGATGCA	TGCACGTTCT	TCAGGGCCTG	AACCTTCTTA	TGCCTTGTCT	2940
	GATAATGAGG	GCTCCCAACA	CATCTTCAC	GTTTCTTACA	GTTCTGCAAT	ACCTGTGCAT	3000
10	GATTCTGTGG	GTGTAACCTA	TCAGGGTTCC	TTATTTAGCG	GCCCTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCACCT	GTTTCTGTAG	CTGAATTTAC	ATATACAACA	3240
	TCTGTGTTTG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGGAAATGAG	3300
	ACTGAACTGC	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
15	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGGCAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACCTTTCAG	TTCAACCTAC	ACATACTGTC	3540
	TCTCAAGCAT	CTGGTGACAC	TTGCTTAAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
20	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATTG	CTACAACCTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAACCTGT	TCTTCCAGCT	GTGCCAGTG	ATCCAATATT	GGTTGAAACC	3780
	CCCAAGTTG	ATAAAATTAG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
	AGTGAAACA	TGCTGCACTC	TACATCTGTA	CCAGTTTTTG	ATGTGTCGCC	TACTTCTCAT	3900
25	ATGCACTCTG	CTTCACTTCA	AGGTTTGACC	ATTTCTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTTTTGTAA	AAAGTGAAAG	TTCCACCAA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAAC	CAGGCCCATC	CCCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATTGATGAA	CCATTAAATA	CACATAATAA	TAAGCTTATA	4140
	CATTCCGATG	AAATTTTAA	CTCCACCAA	AGTTCTGTTA	CTGGTAAGGT	ATTTGCTGGT	4200
30	ATTTCAACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTCTGTTCC	TATAGGAAAT	4260
	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCTGTAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACCTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGACAGA	4440
	GATAGTGATG	GCTTATCCAT	TCATAAGTGT	ATGTCATGCT	CATCCTATAG	AGAATCACAG	4500
35	GAAAAGGTAA	TGAATGATT	AGACACCCAC	GTGCCAGTCT	TTATGGATCA	GAATAATCCA	4560
	ATCTCATACT	CACTATCTGA	GAATTCTGAA	GAAGATAATA	GAGTCACAAG	TGTATCCTCA	4620
	GACAGTCAAA	CTGGTATGGA	CAGAAGTCCT	GGTAAATCAC	CATCAGCAAA	TGGGCTATCC	4680
	CAAAAGCACA	ATGATGGAAA	AGAGGAAAAT	GACATTTCAGA	CTGGTAGTGC	TCTGCTTCCT	4740
	CTCAGCCCTG	AATCTAAAGC	ATGGGCAGTT	CTGACAAGTG	ATGAAGAAAG	TGGATCAGGG	4800
40	CAAGGTACCT	CAGATAGCCT	TAATGAGAAT	GAGACTTCCA	CAGATTTTCC	TTTTGCAGAC	4860
	ACTAATGAAA	AAGATGCTGA	TGGGATCCTG	GCAGCAGGTG	ACTCAGAAAT	AACTCCTGGA	4920
	TTCCCACAGT	CCCCAACATC	ATCTGTTACT	AGCGAGAACT	CAGAAGTGTT	CCACGTTTCA	4980
	GAGGCAGAGG	CCAGTAATAG	TAGCCATGAG	TCTCGTATTG	GTCTAGCTGA	GGGGTTGGAA	5040
	TCCGAGAAGA	AGGCAGTTAT	ACCCCTTGTG	ATCGTGTCTAG	CCCTGACTTT	TATCTGTCTA	5100
45	GTGGTTCTTG	TGGGTATTCT	CATCTACTGG	AGGAAATGCT	TCCAGACTGC	ACACTTTTAC	5160
	TTAGAGGACA	GTACATCCCC	TAGAGTTATA	TCCACACCTC	CAACACCTAT	CTTTCCAATT	5220
	TCAGATGATG	TCGGAGCAAT	TCCAATAAAG	CACCTTCCAA	AGCATGTTGC	AGATTTACAT	5280
	GCAAGTAGTG	GGTTTACTGA	AGAATTTGAG	ACACTGAAAG	AGTTTTACCA	GGAAGTGCAG	5340
	AGCTGTACTG	TTGACTTAGG	TATTACAGCA	GACAGCTCCA	ACCACCCAGA	CAACAAGCAC	5400
50	AAGAATCGAT	ACATAAATAT	CGTTGCCTAT	GATCATAGCA	GGGTAAAGCT	AGCACAGCTT	5460
	GCTGAAAAGG	ATGGCAAACT	GACTGATTAT	ATCAATGCCA	ATTATGTTGA	TGGCTACAAC	5520
	AGACCAAAAG	CTTATATTGC	TGCCCAAGGC	CCACTGAAAT	CCACAGCTGA	AGATTTCTGG	5580
	AGAATGATAT	GGGAACATAA	TGTGGAAGTT	ATTGTCATGA	TAACAAACCT	CGTGGAGAAA	5640
	GGAAGGAGAA	AATGTGATCA	GTACTGGCCT	GCCGATGGGA	GTGAGGAGTA	CGGGAACCTT	5700
55	CTGGTCACTC	AGAAGAGTGT	GCAAGTGCTT	GCCTATTATA	CTGTGAGGAA	TTTTACTCTA	5760
	AGAAACACAA	AAATAAAAAA	GGGCTCCCAG	AAAGGAAGAC	CCAGTGGACG	TGTGGTCACA	5820
	CAGTATCACT	ACACGCAGTG	GCCTGACATG	GGAGTACCAG	AGTACTCCCT	GCCAGTGCTG	5880
	ACCTTTGTGA	GAAAGGCAGC	CTATGCCAAG	CGCCATGCAG	TGGGGCCTGT	TGTCGTCCAC	5940
	TGCAGTGCTG	GAGTTGGAAG	AACAGGCACA	TATATTGTGC	TAGACAGTAT	GTTGCAGCAG	6000
60	ATTCAACACG	AAGGAACCTG	CAACATATTT	GGCTTCTTAA	AACACATCCG	TTCAACAAAG	6060
	AATTATTTGG	TACAAACTGA	GGAGCAATAT	GTCTTCATTC	ATGATACACT	GGTTGAGGCC	6120
	ATACTTAGTA	AAGAACTGA	GGTGCTGGAC	AGTCATATTC	ATGCCTATGT	TAATGCACTC	6180
	CTCATTCCCTG	GACCAGCAGG	CAAAACAAAG	CTAGAGAAAC	AATTCCAGCT	CCTGAGCCAG	6240
	TCAAATATAC	AGCAGAGTGA	CTATTCTGCA	GCCCTAAAGC	AATGCAACAG	GGAAAAGAAT	6300
65	CGAACTTCTT	CTATCATCCC	TGTGGAAAGA	TCAAGGGTTG	GCATTTTATC	CCTGAGTGGA	6360
	GAAAGCAGAG	ACTACATCAA	TGCTTCCCTG	ATCATGGGCT	ATTACCAGAG	CAATGAATTC	6420
	ATCATTACCC	AGCACCTCTT	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	6480
	CATAATGCCC	AACTGGTGGT	TATGATTCCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	6540
	GTTTACTGGC	CAAATAAAGA	TGAGCCTATA	AATTGTGAGA	GCTTTAAGGT	CACTCTTATG	6600
70	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTAA	TAATTCAGGA	CTTTATCTTA	6660
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTCAAGTGTC	TAAATGGCCA	6720
	AATCCAGATA	GCCCCATTAG	TAAACTTATA	GAACCTTATA	GTGTTATAAA	AGAAGAAGCT	6780
	GCCAATAGGG	ATGGGCCTAT	GATTTGTTCT	GATGAGCATG	GAGGAGTGAC	GGCAGGAAC	6840
	TTCTGTGCTC	TGACAACCTT	TATGCACCAA	CTAGAAAAAG	AAAATTCCGT	GGATGTTTAC	6900
75	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	6960
	CAGTTTCTCT	ACAAAGTGAT	CCTCAGCCTT	GTGAGCACAA	GGCAGGAAGA	GAATCCATCC	7020
	ACCTCTCTGG	ACAGTAATGG	TGCAGCATTG	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	7080
	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATGTT	TTTCTCTTTC	7140
	CTAAATTTAG	GCAAGGAAAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCC	ATCACCTGAC	7200
80	AGTAACCTTC	ATGACATAGG	ATTCTGCCGC	CAAATTTATA	TCATTAACAA	TGTGTGCCCT	7260
	TTTGCAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACATAA	ATGATTGAAT	TTTACAGTAT	7320
	TTCTAAGAAT	GGAATTGTGG	TATTTTTTTC	TGTATTGATT	TTAACAGAAA	ATTTCAATTT	7380
	ATAGAGGTTA	GGAATTCCAA	ACTACAGAAA	ATGTTTGTTC	TTAGTGTCAA	ATTTTTAGCT	7440
	GTATTGTAG	CAATTATCAG	GTTTGCTAGA	AATATAACTT	TTAATACAGT	AGCCTGTAAA	7500
	TAAACACTC	TTCATATGTA	TATACACAT	TTTACAACCT	CAGTATTCAC	CTAAAGTAGA	7560
85	AATAATCTGT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	7620
	ATAATTGTAG	ATTTTATAT	TTTACTACTG	AGTCAAGTTT	TCTAGTCTG	TGTAATTGTT	7680
	TAGTTTAATG	ACGTAGTTCA	TTAGCTGGTC	TTACTCTACC	AGTTTCTGTA	CATTGTATTG	7740



TGTTACCTAA GTCATTAACT TTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAAATAGAA 7800  
ATACCTTCAT TTTGAAAGAA GTTTTATGA GAATAACACC TTACCAAACA TTGTTCAAAT 7860  
GGTTTTTATC CAAGGAATTG CAAAAATAAA TATAAATATT GCCATTAAAA AAAAAAAAAA 7920  
AAAAAAAAAA AAAAAAAAAA A

Seq ID NO: 583 Protein sequence  
Protein Accession #: NP\_002842.1

1 11 21 31 41 51  
| | | | | |  
MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK 60  
QSPINIDEDL TQVNVNLKKL KFQGWDKTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
FKASKITFWH GKCNMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEEAV KGKGLRLALS 180  
ILFEVGTEEN LDFKAIIDGV ESVSFRFGKQA ALDPFILLNL LPNSTDKYVI YNGSLTSPPC 240  
15 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVMLMDY LQNNFREQQY KFSRQVFSSY 300  
TGKEEIEHAV CSSEPEENVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360  
HEFLTDGYQD LGAILNLLP NMSYVLQIVA ICTNGLYGKY SDQLIVDMPT DNPDLDFPE 420  
LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTKYNEAKTN 480  
RSPTRGSEFS GKGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPTH VEGTSASLND 540  
20 GSKTVLRSPH MNLSGTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPMEG NVWFPSTDI 660  
TAQPDVGSGR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
TEVTPHAFPT SSRQQDLVST VNVVYSQTTQ PVYNGETPLQ PSYSSEVFPL VTPLLLDNQI 780  
LNTTPAASSS DSALHATPVF PSVDVSFESI LSSYDGAPLL PFSSASFSE LFRHLHTVSQ 840  
25 ILPQVTSATE SDKVPLHASL PVAGGDLLE PSLAQYSDVL STTHAASETL EFGSESGVLY 900  
KTLMFQVEP PSSDAMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG 960  
SLFSGPSHIP IPKSSLITPT ASLLQPTHAL SGDGEWSGAS SDSEFLLPDT DGLTALNISS 1020  
PVSVAEFTYT TSVFGDDNKA LSKSEIIYGN ETELQIPSFN EMVYPSESTV MPNMYDNVKN 1080  
30 LNASLQETSV SISSTKGMFP GSLAHTTTKV FDHEISQVPE NNFVQPTHV VSQASGDTSL 1140  
KPVLSANSEP ASSDPASSEM LSPSTQLLFY ETSASFSTEV LLQPSFQASD VDTLLKTVLP 1200  
AVPSDPILVE TPKVDKISST MLHLIVNSA SSENMLHSTS VPVFDVSPTS HMHSASLQGL 1260  
TISYASEKYE PVLLKSESSH QVVPSTLYSND ELFQTANLEI NQAHPPKGRH VFATPVLSD 1320  
EPLNTLINKL IHSDEILTST KSSVTGKVFA GIPTVASDTF VSTDHSVPIG NGHVAITAVS 1380  
35 PHRDGSVTST KLLFPSKATS ELSHSAKSDA GLVGGGEDGD TDDDGDDDD RDSGLSIHK 1440  
CMSCSSYRES QEKVMNDSST HENSLMDQNN PISYLSSENS EEDNRVTSVS SDSQTGMDRS 1500  
PGKSPSANGS SQKHNDGKEE NDIQTGSALL PLSPEKAWA VLTSDEESGS GQGTSDSLNE 1560  
NETSTDFSFA DTNEKDADGI LAAGDSEITP GFFQSPTSSV TSENSEVFHV SEAEASNSSH 1620  
ESRIGLAEGL ESEKKAIVPL VIVSALTFIC LVVLVGILYI WRKCFQTAHF YLEDSTSPRV 1680  
40 ISTPPTPIFP ISDDVGAIFI KHFPKHVADL HASSGFTEEF ETLKEFYQEV QSTVDLGIT 1740  
ADSSNHPDNK HKNRYINIVA YDHSRVKLAQ LAEKDGKLT D YINANYVDGY NRPKAYIAAQ 1800  
GPLKSTAEDF WRMIWEHNVE VIVMITNLVE KGRRKCDQYW PADGSEYGN FLVTQKSVQV 1860  
LAYYTVRNFY LRNTKIKKGS QKGRPSGRV TQYHYTQWPD MGVPEYSLPV LTFVRKAAYA 1920  
KRHAVGPVVV HCSAGVGRTG TYIVLDSMLQ QIQHEGTVNI FGFLKHRSQ RNYLVQTEEQ 1980  
45 YVFIHDTLVE AILSKETEVL DSHIHAYVNA LLIPGPAGKT KLEKQFQLS QSNIQSDYS 2040  
AALKQCNREK NRTSSIPVE RSRVGISLS EGEGTYINAS YIMGYYSNE FIITQHPLH 2100  
TIKDFWRMIW DHNAQLVMI PDGQNMADDE FVYWPKNDEP INCESFKVTL MAEEHKCLSN 2160  
EEKLIIQDFI LEATQDDYVL EVRHFQCPKW PNPDSPIST FELISVIKEE AANRDGPMIV 2220  
HDEHGGVTAG TFCALTTLMH QLEKENSVDV YQVAKMINLM RPGVFADIEQ YQFLYKVILS 2280  
50 LVSTRQEENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 584 DNA sequence  
Nucleic Acid Accession #: NM\_005688.1  
Coding sequence: 126..4439

1 11 21 31 41 51  
| | | | | |  
CCGGGCGAGT GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCCTGGAGC 60  
AGGGGCGCAG GAATCTGAT GTGAACTAA CAGTCTGTGA GCCCTGGAAC CTCCTGCTCAG 120  
60 AGAAGATGAA GGATATCGAC ATAGGAAAAG AGTATATCAT CCCCAGTCCT GGGTATAGAA 180  
GTGTGAGGGA GAGAACCAGC ACTTCTGGGA CGCACAGAGA CCGTGAAGAT TCCAAGTTCA 240  
GGAGAACTCG ACCGTTGGAA TGCCAAGATG CTTTGGAAAC AGCAGCCCGA GCGGAGGGCC 300  
TCTCTCTTGA TGCCTCCATG CATCTCTCAGC TCAGAATCCT GGATGAGGAG CATCCCAAGG 360  
GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACACCAGC 420  
ACCCAGTGGA CAATGCTGGG CTTTTTCTCT GTATGACTTT TTCGTGGCTT TCTTCTCTGG 480  
65 CCCGTGTGGC CCACAAGAAG GGGGAGCTCT CAATGGAAGA CGTGTGGTCT CTGTCCAAGC 540  
ACGAGTCTTC TGACGTGAAC TGCAGAAGAC TAGAGAGACT GTGGCAAGAA GAGCTGAATG 600  
AAGTTGGGCC AGACGCTGCT TCCCTGCGAA GGGTTGTGTG GATCTTCTGC CGCACCAGGC 660  
TCATCCTGTC CATCGTGTGC CTGATGATCA CGCAGCTGGC TGGCTTCAGT GGACCAGCCT 720  
70 TCATGGTGAA ACACCTCTTG GAGTATACCC AGGCAACAGA GTCTAACCTG CAGTACAGCT 780  
TGTTGTTAGT GCTGGGCCCTC CTCCTGACGG AAATCGTGCG GTCTTGCTCG CTTGCACTGA 840  
CTTGCGCATT GAATTACCGA ACCGCTGTCC GCTTGCGGGG GGCCATCCTA ACCATGGCAT 900  
TTAAGAAGAT CCTTAAGTTA AAGAACATTA AAGAGAAATC CCTGGGTGAG CTCATCAACA 960  
TTTGCTCCAA CGATGGGCAG AGAATGTTTG AGGCAGCAGC CGTTGGCAGC CTGCTGGCTG 1020  
75 GAGGACCCGT TGTTGCCATC TTAGGCATGA TTTATAATGT AATTATTCTG GGACCAACAG 1080  
GCTTCTCTGG ATCAGCTGTT TTTATCCTCT TTTACCCAGC AATGATGTTT GCATCACGGC 1140  
TCACAGCATA TTTACAGGGA AAATGCGTGG CCGCCACGGA TGAACGTGTC CAGAAGATGA 1200  
ATGAAGTTCT TACTTACATT AAATTTATCA AAATGTATGC CTGGGTCAAA GCATTTTCTC 1260  
AGAGTGTTC AAAAAATCCGC GAGGAGGAGC GTCGGATATT GGAAAAAGCC GGGTACTTCC 1320  
AGGGTATCAC TGTGGGTGTG GCTCCCATTT TGGTGGTGAT TGCCAGCGTG GTGACCTTCT 1380  
80 CTGTTTCATAT GACCTTGGGC TTCGATCTGA CAGCAGCACA GGCTTTCACA GTGGTGACAG 1440  
TCTTCAATTC CATGACTTTT GCTTTGAAAG TAACACCGTT TTCAGTAAAG TCCCTCTCAG 1500  
AAGCCTCAGT GGCTGTTGAC AGATTTAAGA GTTTGTTTCT AATGGAAGAG GTTCACATGA 1560  
TAAAGAACAA ACCAGCCAGT CCTCACATCA AGATAGAGAT GAAAAATGCC ACCTTGGCAT 1620  
85 GGGACTCCTC CCACCTCCAGT ATCCAGAACT CGCCCAAGCT GACCCCAAAA ATGAAAAAAG 1680  
ACAAGAGGGC TTCCAGGGGC AAGAAAGAGA AGGTGAGGCA GCTGCAGCGC ACTGAGCATC 1740  
AGGCGGTGCT GGCAGAGCAG AAAGGCCACC TCCTCCTGGA CAGTGACGAG CGGCCAGTC 1800  
CCGAAGAGGA AGAAGGCAAG CACATCCACC TGGGCCACCT GCGCTTACAG AGGACACTGC 1860



	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AACTGGTTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCATT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACCTTC	GCTTATGTGG	CCCAGCAGGC	CTGGATCCTC	AATGCTACTC	2040
5	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTG	CCTGAGGCCT	GACCTGGCCA	TTCTTCCCAG	CAGCGACCTG	ACGGAGATTG	2160
	GAGAGCGAGG	AGCCAACCTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCCGGCCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTT	CAGTGCCTTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
10	TTGTTACCCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGATCTTC	ATGAAAGAGG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACTGATGAA	TTTAAATGGT	GACTATGCTA	2460
	CCATTTTTAA	TAACCTGTTG	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCAGTGG	TTCACAGAAG	AAGTCACAAG	ACAAGGGTCC	TAAAACAGGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
15	GTTTCAGTGC	CTGGTCAGTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
	TCCTGGTTAT	TATGGCCCTT	TTTATGCTGA	ATGTAGGCAG	CACCGCCTTC	AGCACCTGGT	2760
	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAACACCAC	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCCTGA	AAGCCATTCC	AGGAGTTGTC	TTTGTCAAGG	2940
20	GCACGCTGCG	AGCTTCCTCC	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTCGAAGCC	3000
	CTATGAAGTT	TTTTGACACG	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
	TGGATGAAGT	TGACGTGCGG	CTGCCGTTCC	AGGCGAGAT	GTTTCATCCAG	AACGTTATCC	3120
	TGGTGTTCCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCCT	GTGGCAGTGG	3180
	GGCCCCCTGT	CATCCTCTTT	TCAGTCCTGC	ACATTGTCTC	CAGGGTCCTG	ATTCTGGGAGC	3240
25	TGAAGCGTCT	GGACAATATC	ACGCAGTCAC	CTTTCCTCTC	CCACATCACG	TCCAGCATAC	3300
	AGGGCCTTGC	CACCAATCCAC	GCCTACAATA	AAGGGCAGGA	GTTTCTGCAC	AGATAACCAGG	3360
	AGCTGCTGGA	TGACAACCAA	GCTCCTTTTT	TTTTGTTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACACC	CACGGGGCTG	ATGATCGTTC	3480
	TTATGCACGG	GCAGATTCCC	CCAGCTTATG	CGGGTCTCGC	CATCTCTTAT	GCTGTCCAGT	3540
30	TAACGGGGCT	GTTCCAGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTACACT	3600
	CGGTGGAGAG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAGCACCT	GCCAGAATTA	3660
	AGAACAAGGC	TCCCTCCCTT	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGCAG	3720
	AGATGAGGTA	CCGAGAAAAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAAC	3780
	CTAAAGAGAA	GATTGGCATT	GTGGGGCGGA	CAGGATCAGG	GAAGTCCTCG	CTGGGGATGG	3840
35	CCCTCTTCCG	TCTGTGGGAG	TTATCTGAGG	GCTGCATCAA	GATTGATGGA	GTGAGAATCA	3900
	GTGATATTGG	CCTTGCCGAC	CTCCGAAGCA	AACTCTCTAT	CATTCTCTCA	GAGCCGGTGC	3960
	TGTTTCAGTGG	CACTGTGAGA	TCAAATTTGG	ACCCCTTCAA	CCAGTACACT	GAAGACCAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TTGAATCTGA	AGTGATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
40	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
	CCATGGACAC	AGAGACAGAC	TTATTGATTG	AAGAGACCAT	CCGAGAAGCA	TTTGCAGACT	4260
	GTACCATGCT	GACCATTGCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTTCTG	TCCAACGACA	4380
	GTTCCCGATT	CTATGCCATG	TTTGCTGCTG	CAGAGAACAA	GGTCGCTGTC	AAGGGCTGAC	4440
45	TCCTCCCTGT	TGACGAAGTC	TCTTTTCTTT	AGAGCATTCG	CATTCCCTGC	CTGGGGCGGG	4500
	CCCCTCATCG	CGTCTCCTA	CCGAAACCTT	GCCTTTCTCG	ATTTTATCTT	TGCGACAGCA	4560
	GTTCCGGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTGTATTT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTTAA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
50	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
	TATTAAAATA	AGCACTGTGC	TAATAACAGT	GCATATTCCT	TTCTATCATT	TTTGTACAGT	4860
	TTGCTGTACT	AGAGATCTGG	TTTTGTCTAT	AGACTGTAGG	AAGAGTAGCA	TTTCATTCTT	4920
	CTCTAGCTGG	TGTTTTCACG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCCTCTGCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
55	GAGACGGGTG	GGCGGCTGGA	GACCATGCAG	AGCGCCGTGA	GTTCTCAGGG	CTCCTGCCTT	5100
	CTGTCTGGT	GTCACCTACT	GTTTCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCTT	5160
	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCCT	TCTTCTTTT	GCTGTTGTTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTCTT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340
60	GTTGGTTCCA	AGCCCTGGAG	CCAATCTGCG	CTTTTGGAGG	TGGCACTTTT	TCATTTGCCT	5400
	ATTCCACAC	CTCCACAGTT	CAGTGGCAGG	GCTCAGGATT	TCGTGGGTCT	GTTTTCCTTT	5460
	CTCACCGCAG	TCGTCGCACA	GTCTCTCTCT	CTCTCTCCCC	TCAAAGTCTG	CAACTTTAAG	5520
	CAGCTCTTGC	TAATCAGTGT	CTCACACTGG	CGTAGAAGTT	TTTGTACTGT	AAAGAGACCT	5580
	ACCTCAGGTT	GCTGTTGCT	GTGTGGTTTG	GTGTGTTCCC	GCAAAACCCC	TTTGTGCTGT	5640
65	GGGGCTGGTA	GCTCAGGTGG	GCGTGGTCC	TGCTGTCTATC	AGTTGAATGG	TCAGCGTTGC	5700
	ATGTCGTGAC	CAACTAGACA	TTCTGTCCGC	TTAGCATGTT	TGCTGAACAC	CTTGTGGAAG	5760
	CAAAAATCTG	AAAAATGTGA	TAAAATTATT	TTGGATTTTG	TAAAAAATAA	AAAAAATAA	5820
	AAAAAATAA	AAAAAATAA					

Seq ID NO: 585 Protein sequence  
Protein Accession #: NP\_005679.1

	1	11	21	31	41	51	
75	MKDIDIGKEY	IIPSPGYRSV	RERTSTSGTH	RDREDSKFR	TRPLECQDAL	ETAARAEGLS	60
	LDASMHSQRL	ILDEEHPKPK	YHHGLSALKP	IRTTSKHQHP	VDNAGLFSCM	TFSWLSSLAR	120
	VAHKKGELSM	EDVWSLSKHE	SSDVNCRRL	RLWQEELNEV	GPDAASLRRV	VWIFCRTRLI	180
	LSIVCLMITQ	LAGFSGPAPM	VKHLLEYTQA	TESNLQYSL	LVLGLLLTEI	VRWSLALTW	240
	ALNYRTGVRL	RGAILTMAFK	KILKLKNIKE	KSLGELINIC	SNDGQRMFEA	AAVGSLLAGG	300
80	PVVAILGMIY	NVILGPTGF	LGSVAVFILFY	PAMMFASRLT	AYFRKRCVAA	TDERVQKMNE	360
	VLTYIKFIKM	YAWVKAFSOS	VQKIREEERR	ILEKAGYFQG	ITVGVAPIVV	VIASVVTFSV	420
	HMTLGFDLTA	AQAFVTVTVF	NSMTFALKVT	PFSVKSLSEA	SVAVDRFKSL	FLMEVHMIK	480
	NKPASPHIKI	EMKNATLAWD	SSHSSIQNSP	KLTPKMKKDK	RASRGKKEKV	RQLQRTEHQA	540
	VLAEQKHLL	LDSDERPSPE	EEEGKHIHLG	HLRLQRTLHS	IDLEIQEGKL	VGICGSVSGS	600
85	KTSLISAILG	QMTLLEGSIA	ISGTFAYVAQ	QAWILNATLR	DNILFGKEYD	EBRYNSVLNS	660
	CCLRPDLAIL	PSSDLTEIGE	RGANLSGGQR	QRISLARALY	SDRSIYILDD	PLSALDAHVG	720
	NHIFNSAIRK	HLKSKTVLFV	THQLQYLVDC	DEVIFMKEGC	ITERGTHEEL	MNLNGDYATI	780
	FNNLLLGETP	PVEINSKKET	SGSQKKSQDK	GPKTGSVKKE	KAVKPEEGQL	VQLEBKQGS	840



VPWSVYGVYI QAAGGPLAFI VIMALFMLNV GSTAFSTWWL SYWIKQSGSN TTVTRGNETS 900  
VSDSMKDNPH MQYYASIYAL SMAVMLILKA IRGVVFVKGT LRASSRLHDE LFRRILRSPM 960  
KFFDTTPTGR ILNRFSDMD EVDVRLPFQA EMFIQNVILV FFCVGMIAV FPWFLVAVGP 1020  
LVILFSLVHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QEFLHRYQEL 1080  
LDDNQAPFLL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HGQIPPAYAG LAISYAVQLT 1140  
GLFQFTVRLA SETEARFTSV ERINHYIKTL SLEAPARIKN KAPSPDWPQE GEVTFENAEM 1200  
RYRENLPVLV KKVSFYIKPK EKIGIVGRTG SGKSSLGMAL FRLVELSGGC IKIDGVRISD 1260  
IGLADLRSLK SIIPQEPVLF SGTVRSNLDP FNQYTEDQIW DALERTHMKE CIAQLPLKLE 1320  
SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIIQE TIREAFADCT 1380  
MLTIAHRLHT VLGSDDRMVL AQGQVVEFDT PSVLLSNDSS RPYAMFAAAE NKVAVKG

Seq ID NO: 586 DNA sequence  
Nucleic Acid Accession #: NM\_001327.1  
Coding sequence: 89..631

1 11 21 31 41 51  
AGCAGGGGGC GCTGTGTGTA CCGAGAATAC GAGAATACCT CGTGGGGCCCT GACCTTCTCT 60  
CTGAGAGCCG GGCAGAGGCT CCGGAGCCAT GCAGGCCGAA GGCCGGGGCA CAGGGGGTTC 120  
GACGGGCGAT GCTGATGGCC CAGGAGGCCC TGGCATTCCCT GATGGCCAG GGGGCAATGC 180  
TGGCGGGCCA GGAGAGCGCG GTGCCACGGG CGGCAGAGGT CCCCAGGGCG CAGGGGCAGC 240  
AAGGGCCCTG GGGCCGGGAG GAGGCGCCCC GCGGGGTCCG CATGGCGGCG CGGCTTCAGG 300  
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGGAG AGCCGCCTGC TTGAGTTCTA 360  
CCTCGCCATG CCTTTCGCGA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420  
GGATGCCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACTG TGTCCGGCAA 480  
CATACTGACT ATCCGACTGA CTGCTGCAGA CCACCGCCAA CTGCAGCTCT CCATCAGCTC 540  
CTGTCTCCAG CAGCTTTCCT TGTGATGTG GATCACGCGT TGCTTCTGCG CCGTGTCTTT 600  
GGCTCAGCCT CCTCAGGGC AGAGGCGCTA AGCCAGCCTT GCGCCCCCTT CCTAGGTCAT 660  
GCCTCCTCCC CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720  
GTTTGTGCTG GGAGGAGGAC GGCTTACATG TTTGTTTCTG TAGAAAATAA AACTGAGCTA

Seq ID NO: 587 Protein sequence  
Protein Accession #: NP\_001318.1

1 11 21 31 41 51  
MQAEGRGTTG STGDADGPGG PGIPDGPNGN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60  
PRGPHGGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EAELARRSLA QDAPPLPVP 120  
VLLKEFTVSG NLTIRLTAA DHRQLQLSIS SCLQLSLLM WITQCFLPVF LAQPPSGQRR

Seq ID NO: 588 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 52..459

1 11 21 31 41 51  
CCTCGTGGGC CCTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGGAGC CATGCAGGCC 60  
GAAGGCCAGG GCACAGGGGG TTCGACGGGC GATGCTGATG GCCCAGGAGG CCCTGGCATT 120  
CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CGGGTGCCAC GGGCGGCAGA 180  
GGTCCCCGGG GCGCAGGGGC AGCAAGGGCC TCGGGGCCGA GAGGAGGCGC CCCGCGGGGT 240  
CCGCATGGCG GTGCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTGCGGGGC CAGGAGGCCG 300  
GACAGCCGCT TGCTTCAGTT CCGACTGACT GCTGCAGACC ACCGCCAACT GCAGCTCTCC 360  
ATCAGCTCCT GTCTCCAGCA GCTTTCCTTG TTGATGTGGA TCACGAGTGC CTTTCTGCCC 420  
GTGTTTTTGG CTCAGGCTCC CTCAGGGCAG AGGCGCTAAG CCCAGCCTGG CGCCCTTCC 480  
TAGGTCATGC CTCCTCCCCT AGGGAATGGT CCCAGCACGA GTGGCCAGTT CATTGTGGGG 540  
GCCTGATTGT TTGTCGCTGG AGGAGGACGG CTTACATGTT TGTTTCTGTA GAAAATAAAG 600  
CTGAGCTA

Seq ID NO: 589 Protein sequence  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
MQAEGQGTGG STGDADGPGG PGIPDGPNGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60  
PRGPHGGAAS AQDGRCPGCA RRPDSRLLQF RLTAADHRQL QLSISSCLQQ LSLLMWITQC 120  
FLPVFLAQAP SGQRR

Seq ID NO: 590 DNA sequence  
Nucleic Acid Accession #: NM\_005562.1  
Coding sequence: 90..3671

1 11 21 31 41 51  
ACAGCGGAGC GCAGAGTGAG AACCACCAAC CGAGGCGCCG GGCAGCGACC CCTGCAGCGG 60  
AGACAGAGAC TGAGCGGCCC GGCACCGCCA TGCCTGCGCT CTGGCTGGGC TGCTGCCTCT 120  
GCTTCTCGCT CCTCTGCCC GCAGCCCGGG CCACCTCCAG GAGGGAAGTC TGTGATTGCA 180  
ATGGGAAGTC CAGGCAGTGT ATCTTTGATC GGAACAATTCA CAGACAACT GGTAATGGAT 240  
TCCGCTGCCT CAACTGCAAT GACAACACTG ATGGCAATTCA CTGCGAGAAG TGCAAGAATG 300  
GCTTTTACCG GCACAGAGAA AGGGACCGCT GTTTGCCCTG CAATTGTAAC TCCAAAGGTT 360  
CTCTTAGTGC TCGATGTGAC AACTCTGGAC GGTGCAGCTG TAAACCAGGT GTGACAGGAG 420  
CCAGATGCGA CCGATGTCTG CCAGGCTTCC ACATGCTCAC GGATGCGGGG TGCACCCAAG 480  
ACCAGAGACT GCTAGACTCC AAGTGTGACT GTGACCCAGC TGGCATCGCA GGGCCCTGTG 540  
ACGCGGGCCG CTGTGTCTGC AAGCCAGACT TTAGTGGAGA ACGCTGTGAT AGGTGTCGAT 600  
CAGGTTACTA TAATCTGGAT GGGGGGAACC CTAGGGGCTG TACCCAGTGT TTCTGCTATG 660  
GGCATTGAGC CAGCTGCCGC AGCTCTGCAG AATACAGTGT CCATAAGATC ACCTCTACCT 720  
TTCATCAAGA TGTGATGGC TGAAGGCTG TCCAACGAAA TGGGTCTCCT GCAAAGCTCC 780  
AATGGTCACA GCGCCATCAA GATGTGTTTA GCTCAGCCCA ACGACTAGAC CCTGTCTATT 840



	TTGTGGCTCC	TGCCAAATTT	CTTGGGAATC	AACAGGTGAG	CTATGGGCAA	AGCCTGTCCT	900
	TTGACTACCG	TGTGGACAGA	GGAGGCAGAC	ACCCATCTGC	CCATGATGTG	ATTCTGGAAG	960
	GTGCTGGTCT	ACGGATCACA	GCTCCCTTGA	TGCCACTTGG	CAAGACACTG	CCTTGTGGGC	1020
5	TCACCAAGAC	TTACACATTC	AGGTTAAATG	AGCATCCAAG	CAATAATTGG	AGCCCCCAGC	1080
	TGAGTTACTT	TGAGTATCGA	AGGTTACTGC	GGAATCTCAC	AGCCCTCCGC	ATCCGAGCTA	1140
	CATATGGAGA	ATACAGTACT	GGGTACATTG	ACAATGTGAC	CCTGATTTCA	GCCCGCCCTG	1200
	TCTCTGGAGC	CCCAGCACCC	TGGGTTGAAC	AGTGTATATG	TCTGTGTTGG	TACAAGGGGC	1260
	AATTCTGCCA	GGATTGTGCT	TCTGGCTACA	AGAGAGATTC	AGCGAGACTG	GGGCCTTTTG	1320
10	GCACCTGTAT	TCCTTGTAAC	TGTCAAGGGG	GAGGGGCTTG	TGATCCAGAC	ACAGGAGATT	1380
	GTTATTACAG	GGATGAGAAT	CCTGACATTG	AGTGTGCTGA	CTGCCCAATT	GGTTTCTACA	1440
	ACGATCCGCA	CGACCCCCGC	AGCTGCAAGC	CATGTCCCTG	TCATAACGGG	TTCAGCTGCT	1500
	CAGTGATGCC	GGAGACGGAG	GAGGTGGTGT	GCAATAACTG	CCCTCCCGGG	GTCACCGGTG	1560
	CCCGCTGTGA	GCTCTGTGCT	GATGGCTACT	TTGGGGACCC	CTTTGGTGAA	CATGGCCCAG	1620
	TGAGGCCCTG	TCAGCCCTGT	CAATGCAACA	ACAATGTGGA	CCCCAGTGCC	TCTGGGAATT	1680
15	GTGACCGGCT	GACAGGCAGG	TGTTTGAAGT	GTATCCACAA	CACAGCCGGC	ATCTACTGCG	1740
	ACCAGTGCAA	AGCAGGCTAC	TTCGGGGACC	CATTGGCTCC	CAACCCAGCA	GACAAGTGTC	1800
	GAGCTTGCAA	CTGTAACCCC	ATGGGCTCAG	AGCCTGTAGG	ATGTCGAAGT	GATGGCACCT	1860
	GTGTTTGCAA	GCCAGGATTT	GGTGGCCCCA	ACTGTGAGCA	TGGAGCATTG	AGCTGTCCAG	1920
20	CTTGCTATAA	TCAAGTGAAAG	ATTGAGATGG	ATCAGTTTAT	GCAGCAGCTT	CAGAGAATGG	1980
	AGGCCCTGAT	TTCAAAGGCT	CAGGGTGGTG	ATGGAGTAGT	ACCTGATACA	GAGCTGGAAG	2040
	GCAGGATGCA	GCAGGCTGAG	CAGGCCCTTC	AGGACATTCT	GAGAGATGCC	CAGATTTTCA	2100
	AAGGTGCTAG	CAGATCCCTT	GGTCTCCAGT	TGGCCAAGGT	GAGGAGCCAA	GAGAACAGCT	2160
	ACCAGAGCCG	CCTGGATGAC	CTCAAGATGA	CTGTGGAAAG	AGTTCGGGCT	CTGGGAAGTC	2220
25	AGTACCAGAA	CCGAGTTCGG	GATACTCACA	GGCTCATCAC	TCAGATGCAG	CTGAGCCTGG	2280
	CAGAAAGTGA	AGCTTCCCTG	GGAAACACTA	ACATTCCCTG	CTCAGACCAC	TACGTGGGGC	2340
	CAAATGGCTT	TAAAAGTCTG	GCTCAGGAGG	CCACAAGATT	AGCAGAAAGC	CACGTTGAGT	2400
	CAGCCAGTAA	CATGGAGCAA	CTGACAAGGG	AAACTGAGGA	CTATTCCAAA	CAAGCCCTCT	2460
	CAGTGGTGCG	CAAGGCCCTG	CATGAAGGAG	TCGGAAGCGG	AAGCGGTAGC	CCGGACGGTG	2520
30	CTGTGGTGCA	AGGGCTTGTC	GAAAAATTGG	AGAAAACCAA	GTCCCTGGCC	CAGCAGTTGA	2580
	CAAGGGAGGC	CACTCAAGCG	GAAATTGAAG	CAGATAGGTC	TTATCAGCAC	AGTCTCCGCC	2640
	TCCTGGATTG	AGTGTCTCGG	CTTCAGGGAG	TCAGTGATCA	GTCCCTTCAG	GTGGAAGAAG	2700
	CAAAGAGGAT	CAAAACAAAA	GCGGATTAC	TCTCAACGCT	GGTAACCAGG	CATATGGATG	2760
	AGTTCAAGCG	TACACAAAAG	AATCTGGGAA	ACTGGAAAGA	AGAAGCACAG	CAGCTCTTAC	2820
35	AGAATGGAAA	AAGTGGGAGA	GAGAAATCAG	ATCAGCTGCT	TTCCCGTGCC	AATCTTGCTA	2880
	AAAGCAGAGC	ACAAGAAGCA	CTGAGTATGG	GCAATGCCAC	TTTTTATGAA	GTTGAGAGCA	2940
	TCCTTAAAAA	CCTCAGAGAG	TTTGACCTGC	AGGTGGACAA	CAGAAAAGCA	GAAGCTGAAG	3000
	AAGCCATGAA	GAGACTCTCC	TACATCAGCC	AGAAGGTTTC	AGATGCCAGT	GACAAGACCC	3060
	AGCAAGCAGA	AAGAGCCCTG	GGGAGCGCTG	CTGCTGATGC	ACAGAGGGCA	AAGAATGGGG	3120
40	CCGGGGAGGC	CCTGGAATC	TCCAGTGAGA	TTGAACAGGA	GATTGGGAGT	CTGAACTTGG	3180
	AAGCCAATGT	GACAGCAGAT	GGAGCCTTGG	CCATGGAAAA	GGGACTGGCC	TCTCTGAAGA	3240
	GTGAGATGAG	GGAAAGTGAA	GGAGAGCTGG	AAAGGAAGGA	GCTGGAGTTT	GACACGAATA	3300
	TGGATGCAGT	ACAGATGGTG	ATTACAGAAG	CCCAGAAGGT	TGATACCAGA	GCCAAGAACG	3360
	CTGGGGTTAC	AATCCAAGAC	ACACTCAACA	CATTAGACGG	CCTCCTGCAT	CTGATGGACC	3420
45	AGCCTCTCAG	TGTAGATGAA	GAGGGGCTGG	TCTTACTGGA	GCAGAAGCTT	TCCCAGGCCA	3480
	AGACCCAGAT	CAACAGCCAA	CTGCGGCCCC	TGATGTCAGA	GCTGGAAGAG	AGGGCACGTC	3540
	AGCAGAGGGG	CCACCTCCAT	TTGCTGGAGA	CAAGCATAGA	TGGGATTCTG	GCTGATGTGA	3600
	AGAACTTGGA	GAAACCTTAG	GACAACCTGC	CCCCAGGCTG	CTACAATACC	CAGGCTCTTG	3660
	AGCAACAGTG	AAGCTGCCAT	AAATATTTCT	CAACTGAGGT	TCTTGGGATA	CAGATCTCAG	3720
50	GGCTCGGGAG	CCATGTCATG	TGAGTGGGTG	GGATGGGGAC	ATTGGAACAT	GTTTAATGGG	3780
	TATGCTCAGG	TCAACTGACC	TGACCCCAT	CCTGATCCCA	TGGCCAGGTG	GTTGTCTTAT	3840
	TGCACCATAC	TCCTTGCTTC	CTGATGCTGG	GCAATGAGGC	AGATAGCACT	GGGTGTGAGA	3900
	ATGATCAAGG	ATCTGGACCC	CAAAGAATAG	ACTGGATGGA	AAGACAAACT	GCACAGGCAG	3960
	ATGTTTGCCCT	CATAATAGTC	GTAAGTGGAG	TCCTGGAATT	TGGACAAAGT	CTGTTGGGAT	4020
55	ATAGTCAACT	TATTCCTTGA	GTAATGTGAC	TAAAGGAAAA	AACTTTGACT	TTGCCCAGGC	4080
	ATGAAATTCT	TCCTAATGTC	AGAACAGAGT	GCAACCCAGT	CACACTGTGG	CCAGTAAAT	4140
	ACTATTGCCT	CATATTGTCC	TCTGCAAGCT	TCTTGCTGAT	CAGAGTTCCCT	CCTACTTACA	4200
	ACCCAGGGTG	TGAACATGTT	CTCCATTTTC	AAGCTGGAAG	AAGTGAGCAG	TGTTGGAGTG	4260
	AGGACCTGTA	AGGCAGGCCC	ATTCAGAGCT	ATGGTGCTTG	CTGGTGCCCTG	CCACCTTCAA	4320
60	GTTCCTGGAC	TGGGCATGAC	ATCCTTTCTT	TTAATGATGC	CATGGCAACT	TAGAGATTGC	4380
	ATTTTATTAT	AAGCATTTCC	TACCAGCAAA	GCAAATGTTG	GGAAAGTATT	TACTTTTTCG	4440
	GTTTCAAAGT	GATAGAAAAG	TGTGGCTTGG	GCATTGAAAG	AGGTAAAATT	CTCTAGATTT	4500
	ATTAGTCCTA	ATTCAATCCT	ACTTTTCGAA	CACCAAAAAT	GATGCGCATC	AATGTATTTT	4560
	ATCTTATTTT	CTCAATCTCC	TCTCTCTTTC	CTCCACCCAT	AATAAGAGAA	TGTTCCCTACT	4620
65	CACACTTCAG	CTGGGTCACA	TCCATCCCTC	CATTATCCCT	TCCATCCATC	TTTCCATCCA	4680
	TTACCTCCAT	CCATCCTTCC	AACATATATT	TATTGAGTAC	CTACTGTGTG	CCAGGGGCTG	4740
	GTGGGACAGT	GGTGACATAG	TCTTGCCCTT	CATAGAGTTG	ATTGTCTAGT	GAGGAAGACA	4800
	AGCATTTTTA	AAAAATAAAT	TTAAACTTAC	AAACTTTGTT	TGTCACAAGT	GGTGTATTAT	4860
	GCAATAACCG	CTTGTTTTCG	AACCTCTTTG	CTCAACAGAA	CATATGTTGC	AAGACCCTCC	4920
70	CATGGGGGCA	CTTGAGTTT	GGCAAGGCTG	ACAGAGCTCT	GGGTTGTGCA	CATTTCTTTG	4980
	CATTCCAGCT	GTCACCTCTG	GCCTTCTTAC	AACTGATTGC	AACAGACTGT	TGAGTTATGA	5040
	TAACACCAGT	GGGAATTGCT	GGAGGAACCA	GAGGCACTTC	CACCTTGGCT	GGGAAGACTA	5100
	TGGTGCTGCC	TTGCTTCTGT	ATTTCTTGG	ATTTTCTGTA	AAGTGTTTTT	AAATAAAGAA	5160
	CAATTGTTAG	ATGCC					

Seq ID NO: 591 Protein sequence  
Protein Accession #: NP\_005553.1

	1	11	21	31	41	51	
80	MPALWLGCCL	CFSLLLPAAR	ATSRREVDCD	NGKSRQCIFD	RELHRQTGNG	FRCLNCNDNT	60
	DGIHCEKCKN	GFYRHRERDR	CLPCNNSKSG	SLSARCDNSG	RCCKPGVTG	ARCDRLPGF	120
	HMLTDAGCTQ	DQRLDLSKCD	CDPAGIAGPC	DAGRCVCKPA	VTGERCDRCR	SGYYNLDGGN	180
	PEGCTQCFCY	GHSASCRSSA	EYSVHKITST	FHQDVGWKA	VQRNGSPAKL	QWSQRHQDVF	240
	SSAQRLLDPVY	TVAPAKFLGN	QQVSYGQSLG	FDYRVDGRGR	HPSAHDVILE	GAGLRITAPL	300
85	MPLGKTLPCG	LTKTYTFRLN	EHPSNNWSPQ	LSYFEYRRL	RNLTLRIRA	TYGEYSTGYI	360
	DNVTILISARP	VSGAPAPWVE	QCICPVGYKG	QFCQDCASGY	KRDSARLGPF	GTCIPCNCQG	420
	GGACDPDTGD	CYSGDENPDI	ECADCPIGFY	NDPHDPRSC	PCPCHNGFSC	SVMPEETEEV	480



CNNCPPPGVTG ARCELCADGY FGDPFGEHGP VRPCQPCQCN NNVDPSASGN CDRLTGRCCLK 540  
 CIHNTAGIYC DQCKAGYFGD PLAPNPADK RACNCNPMGS EPVGCRSDGT CVCKPGFGGP 600  
 NCEHGAFCSP ACYNQVKIQM DQFMQQLQRM EALISKAQGG DGVVPDTELE GRMQQAEQAL 660  
 QDILRDAQIS EGASRSLGLQ LAKVRSQENS YQSRLDDLKM TVERVRALGS QYQNRVRDTH 720  
 RLITQMQLSL AESEASLGNT NIPASDHYVG PNGFKSLAQE ATRLAESHVE SASNMEQLTR 780  
 ETEDYSKQAL SLVRKALHEG VSGSGSGSPDG AVVQGLVEKL ETKSLAQQL TREATQAEIE 840  
 ADRSYQHSLR LLDSVSRLOQ VSDQSFQVEE AKRIKQKADS LSTLVTRHMD EFKRTQKNLG 900  
 NWKEEAQQLL QNGKSGREKS DQLLSRANLA KSRAQEALSM GNATFYEVES ILKNLREFDL 960  
 QVDNRKAEAE EAMKRLSYIS QKVSDASDKT QQAERALGSA AADAQRAKNG AGEALEISSE 1020  
 IEQEIGSLNL EANVTADGAL AMEKGLASLK SEMREVEGEL ERKELEFDTN MDAVQMVITE 1080  
 AQKVDTRAKN AGVTIQDTLN TLDGLLHLM D QPLSVDEEGL VLLEQKLSRA KTQINSQLRP 1140  
 MMSELEERAR QQRGHLHLE TSIDGILADV KNLENIRDNL PPGCYNTQAL EQQ

Seq ID NO: 592 DNA sequence  
 Nucleic Acid Accession #: AF101051.1  
 Coding sequence: 221.856

1 11 21 31 41 51  
 | | | | | |  
 20 GAGCAACCTC AGCTTCTAGT ATCCAGACTC CAGCGCCGCC CCGGGCGCGG ACCCCAACCC 60  
 CGACCCAGAG CTTCTCCAGC GGCGCGCGAG CGAGCAGGGC TCCCCGCCTT AACTTCCTCC 120  
 GCGGGGCCCA GCCACCTTCG GGAGTCCGGG TTGCCACCT GCAAACCTCTC CGCCTTCTGC 180  
 ACCTGCCACC CCTGAGCCAG CGCGGGCGCC CGAGCGAGTC ATGGCCAACG CGGGGCTGCA 240  
 25 GCTGTTGGGC TTCATTCTCG CCTTCTGGG ATGGATCGGC GCCATCGTCA GCACTGCCCT 300  
 GCCCCAGTGG AGGATTTACT CCTATGCCGG CGACAACATC GTGACCGCCC AGGCCATGTA 360  
 CGAGGGGCTG TGGATGTCTT GCGTGTGCGA GAGCACCAGG CAGATCCAGT GCAAAGTCTT 420  
 TGACTCCTTG CTGAATCTGA GCAGCAACAT GCAGCAACC CGTGCCTTGA TGGTGGTTGG 480  
 CATCCTCCTG GGAGTGATAG CAATCTTTGT GGCCACCGTT GGCATGAAGT GTATGAAGTG 540  
 30 CTTGGAAGAC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTTCT 600  
 TCTTGCAAGT CTGGCTATTT TAGTTGCCAC AGCATGGTAT GGCAATAGAA TCGTTCAAGA 660  
 ATTCTATGAC CCTATGACCC CAGTCAATGC CAGGTACGAA TTGGTTCAGG CTCTCTTCAC 720  
 TGGCTGGGCT GCTGCTTCTC TCTGCCCTCT GGGAGGTGCC CTACTTTGCT GTTCCTGTCC 780  
 CCGAAAAACA ACCTCTTACC CAACACCAAG GCCCTATCCA AAACCTGCAC CTTCCAGCGG 840  
 35 GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAAAATCA TGTGAAACA AACCGAAAAAT 900  
 GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GTAATCTGAA 960  
 GTATGGTATT AAAAAACAAA CAAACAAACA AAAAACCCTAT GTGTTAAAAAT ACTCAGTGCT 1020  
 AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080  
 TTGTATTACT GCTTCCCAT T GAGTAATCAT ACTCAAATGG GGGAAGGGGT GCTCCTTAAA 1140  
 40 TATATATAGA TATGTATATA TACATGTTTT TCTATTAAAA ATAGACAGTA AAATACTATT 1200  
 CTCATTATGT TGATACTAGC ATACTTAAAA TATCTTAAAA ATAGGTAAAT GTATTTAATT 1260  
 CCATATTGAT GAAGATGTTT ATTGGTATAT TTTCTTTTTT GTCCTTATAT ACATATGTAA 1320  
 CAGTCAAATA TCATTTACTC TTCTTCATTA GCTTTGGGTG CCTTTGCCAC AAGACCTAGC 1380  
 CTAATTTACC AAGGATGAAT TCTTCAATT CTTTCATGCGT GCCCTTTTCA TATACTTATT 1440  
 TTATTTTTTA CCATAATCTT ATAGCACTTG CATCGTTATT AAGCCCTTAT TTGTTTTGTG 1500  
 45 TTTCAATTGGT CTCTATCTCC TGAATCTAAC ACATTTTATA GCCTACATTT TAGTTTCTAA 1560  
 AGCCAAGAAG AATTTATTAC AAATCAGAAC TTTGGAGGCA AATCTTTCTG CATGACCAA 1620  
 GTGATAAATT CCTGTTGACC TTCCACACACA ATCCCTGTAC TCTGACCCAT AGCACTCTTG 1680  
 TTTGCTTTGA AAATATTTGT CCAATTGAGT AGCTGCATGC TGTTCACCCA GGTGTTGTAA 1740  
 50 CACAACCTTA TTGATTGAAT TTTTAAGCTA CTTATTCTAT GTTTTATATC CCCCTAAACT 1800  
 ACCTTTTTGT TCCCCATTCC TTAATTGTAT TGTTCCTCCA AGTGTAAATTA TCATGCGTTT 1860  
 TATATCTTCC TAATAAGGTG TGGTCTGTTT GTCTGAACAA AGTGCTAGAC TTTCTGGAGT 1920  
 GATAATCTGG TGACAAATAT TCTCTCTGTA GCTGTAAGCA AGTCACCTAA TCTTTCTACC 1980  
 TCTTTTTTCT ATCTGCCAAA TTGAGATAAT GATACTTAAC CAGTTAGAAG AGGTAGTGTG 2040  
 55 AATATTAATT AGTTTATATT ACTCTCATTC TTTGAACATG AACTATGCCT ATGTAGTGTC 2100  
 TTTATTTGCT CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACCT ACACACGTAC 2160  
 CTTTCATGTA TTCCTGCTCT TCTCTCTCTT ACCAGTCTAT TTCCTACTGAA CAAAACCTAC 2220  
 ACACATACCT TCATGTGGTT CAGTGCCCTC CTCTCTCTAC CAGTCTATTT CCACTGAACA 2280  
 AAACCTACCG ACATACCTTC ATGTGGCTCA GTGCCCTTCT CTCTCTACCA GTCTATTTCC 2340  
 60 ATTCTTTTCA CTGTGTCTGA CATGTTTCTG TCTGTTTCCA TTTTAACAAC TGCTCTTACT 2400  
 TTTCCAGTCT GTACAGAATG CTATTTCACT TGAGCAAGAT GATGTATGGA AAGGGTGTG 2460  
 GCACTGGTGT CTGGAGACCT GGATTTGAGT CTTGGTGCTA TCAATCACCG TCTGTGTTG 2520  
 AGCAAGGCAT TTGGCTGCTG TAAGCTTATT GCTTCATCTG TAAGCGGTG TTTGTAATTC 2580  
 CTGATCTTCC CACCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640  
 65 GTGGTTTTGT AATTTGAAAA GTGCTATACT AAGGGAAGA ATTGAGGAAT TAACTGCATA 2700  
 CGTTTTGGTG TTGCTTTTCA AATGTTTGAA AATAAAAAAA TGTAAAGAAA TGGGTTTCTT 2760  
 GCCTTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAAACGAAT 2820  
 AAGATTCTGA GGAAGTCTTA TCTTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880  
 ACAGATGTAA TGGGAAGAAA TAAAAGCCTA CGTGTGTTGTA AATCCAACAG CAAGGGAGAT 2940  
 70 TTTTGAATCA TAATAACTCA TAAGGTGCTA TCTGTTTCACT GATGCCCTCA GAGCTCTTGC 3000  
 TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGGTACT TCATAATAAA 3060  
 CTACACAAGG AAAGTCAGCC ACCGTGCTTT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120  
 TGCCTTCCAA ACCTGAGAAT ATATGCTTTT GGAAGTTAAA ATTTAAATGG CTTTGGCCAC 3180  
 ATACATAGAT CTTTCATGAT TGTGAGTGTA ATTCCATGTG GATATCAGTT ACCAAACATT 3240  
 75 AAAAAAAAT TTTATGGCCC AAAATGACCA ACGAAATTGT TACAATAGAA TTTATCCAAT 3300  
 TTTGATCTTT TTATATTCTT CTACCAACAC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360  
 TTATAATGGG AATTTGTATA AAGCATTACT CTTTTTCAAT AAATTGTTTT TTAATTTAAA 3420  
 AAAAGGAAAA AAAAAAAAAA AAA

Seq ID NO: 593 Protein sequence  
 Protein Accession #: AAD16433.1

1 11 21 31 41 51  
 | | | | | |  
 85 MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMYEGL WMSCVSQSTG 60  
 QIQCKVFDSL LNLSTLQAT RALMVVGILL GVIAIFVATV GMKCMKCLD DEVQKMRMAV 120  
 IGGAIFFLAG LAILVATAWY GNRIVQEFYD PMTPVNNARYE FGQALFTGWA AASLCLLGG 180  
 LLCCSCPRKT TSYPTPRPYP KPAESSGKDY V



Seq ID NO: 594 DNA sequence  
Nucleic Acid Accession #: NM\_006180.1  
Coding sequence: 352..2820

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70

1	11	21	31	41	51	
CCCCCATTCG	CATCTAACAA	GGAATCTGCG	CCCCAGAGAG	TCCCGGACGC	CGCCGGTCGG	60
TGCCCCGGCG	GCCGGGCCAT	GCAGCGACGG	CCGCCCCGGA	GCTCCGAGCA	GCGGTAGCGC	120
CCCCCTGTAA	AGCGGTTTCG	TATGCCGGGA	CCACTGTGAA	CCCTGCCGCC	TGCCGGAACA	180
CTCTTCGCTC	CGGACCAGCT	CAGCCTCTGA	TAAGCTGGAC	TCGGCACGCC	CGCAACAAGC	240
ACCGAGGAGT	TAAGAGAGCC	GCAAGCGCAG	GGAAGGCCTC	CCCGCACGGG	TGGGGGAAAG	300
CGGCCGGTGC	AGCGCGGGGA	CAGGCACTCG	GGCTGGCACT	GGCTGCTAGG	GATGTCGTCC	360
TGGATAAGGT	GGCATGGACC	CGCCATGGCG	CGGCTCTGGG	GCTTCTGCTG	GCTGGTTGTG	420
GGCTTCTGGA	GGGCGGCTTT	CGCCTGTCCC	ACGTCCTGCA	AATGCAGTGC	CTCTCGGATC	480
TGGTGCAGCG	ACCCCTTCTCC	TGGCATCGTG	GCATTTCCGA	GATTGGAGCC	TAACAGTGTA	540
GATCCTGAGA	ACATCACCGA	AATTTTCATC	GCAAACCAAG	AAAGGTTAGA	AATCATCAAC	600
GAAGATGATG	TTGAAGCTTA	TGTGGGACTG	AGAAATCTGA	CAATTGTGGA	TTCTGGATTA	660
AAATTTGTGG	CTCATAAAGC	ATTTCGTAAA	AACAGCAACC	TGCAGCACAT	CAATTTTACC	720
CGAAACAAAC	TGACGAGTTT	GTCTAGGAAA	CATTTCCGTC	ACCTTGACTT	GTCTGAACTG	780
ATCCTGGTGG	GCAATCCATT	TACATGCTCC	TGTGACATTA	TGTGGATCAA	GACTCTCCAA	840
GAGGCTAAAT	CCAGTCCAGA	CACTCAGGAT	TTGTACTGCC	TGAATGAAAG	CAGCAAGAAT	900
ATTCCCCTGG	CAAACCTGCA	GATACCCAAT	TGTGGTTTGC	CATCTGCAAA	TCTGGCCGCA	960
CCTAACCTCA	CTGTGGAGGA	AGGAAAGTCT	ATCACATTAT	CCTGTAGTGT	GGCAGGTGAT	1020
CCGGTTCCTA	ATATGTATTG	GGATGTTGGT	AACCTGGTTT	CCAAACATAT	GAATGAAACA	1080
AGCCACACAC	AGGGCTCCTT	AAGGATAACT	AACATTTTAT	CCGATGACAG	TGGGAAGCAG	1140
ATCTCTTGTG	TGGCGGAAAA	TCTTGTAGGA	GAAGATCAAG	ATTCTGTCAA	CCTCACTGTG	1200
CATTTTGCAC	CAACTATCAC	ATTTCTCGAA	TCTCCAACCT	CAGACCACCA	CTGGTGCATT	1260
CCATTCACTG	TGAAAGGCAA	CCCCAAACCA	GCGCTTCAGT	GGTTCATATA	CGGGGCAATA	1320
TTGAATGAGT	CCAAATACAT	CTGTACTAAA	ATACATGTTA	CCAATCACAC	GGAGTACCAC	1380
GGCTGCCTCC	AGCTGGATAA	TCCCACTCAC	ATGAACAATG	GGGACTACAC	TCTAATAGCC	1440
AAGAATGAGT	ATGGGAAGGA	TGAGAAACAG	ATTTCTGCTC	ACTTCATGGG	CTGGCCTGGA	1500
ATTGACGATG	GTGCAAACCC	AAATTATCCT	GATGTAATTT	ATGAAGATTA	TGGAAGTACA	1560
GCGAATGACA	TGCGGACAC	CACGAACAGA	AGTAATGAAA	TCCCTTCCAC	AGACGTCATC	1620
GATAAAACCG	GTCGGGAACA	TCTCTCGGTC	TATGCTGTGG	TGGTGATTGC	GTCTGTGGTG	1680
GGATTTTGCC	TTTTGGTAAT	GCTGTTTCTG	CTTAAGTTGG	CAAGACACTC	CAAGTTTGGC	1740
ATGAAAGGCC	CAGCCTCCGT	TATCAGCAAT	GATGATGACT	CTGCCAGCCC	ACTCCATCAC	1800
ATCTCCAATG	GGAGTAACAC	TCCATCTTCT	TCGGAAGGTG	GCCCAGATGC	TGTCATTATT	1860
GGAATGACCA	AGATCCCTGT	CATTGAAAAAT	CCCCAGTACT	TTGGCATCAC	CAACAGTCAG	1920
CTCAAGCCAG	ACACATTTGT	TCAGCACATC	AAGCGACATA	ACATTGTTCT	GAAAAGGGAG	1980
CTAGGCGAAG	GAGCCTTTGG	AAAAGTGTTC	CTAGCTGAAT	GCTATAACCT	CTGTCCTGAG	2040
CAGGACAAGA	TCTTGGTGGC	AGTGAAGACC	CTGAAGGATG	CCAGTGACAA	TGCACGCAAG	2100
GACTTCCACC	GTGAGGCCGA	GCTCCTGACC	AACCTCCAGC	ATGAGCACAT	CGTCAAGTTC	2160
TATGGCGTCT	GCGTGGAGGG	CGACCCCTC	ATCATGGTCT	TTGAGTACAT	GAAGCATGGG	2220
GACCTCAACA	AGTTCTCTAG	GGCACACGGC	CCTGATGCCG	TGCTGATGGC	TGAGGGCAAC	2280
CCGCCCACGG	AACCTGACGA	GTGCGAGATG	CTGACATAG	CCCAGCAGAT	CGCCGCGGGC	2340
ATGGTCTACC	TGGCGTCCCA	GCACTTCGTC	CACCGCGATT	TGGCCACCAG	GAAGTGCCTG	2400
GTCGGGGAGA	ACTTGCTGGT	GAAAATCGGG	GACTTTGGGA	TGTCCCGGGA	CGTGTACAGC	2460
ACTGACTACT	ACAGGGTCGG	TGGCCACACA	ATGCTGCCCC	TTGCTGGGAT	GCCTCCAGAG	2520
AGCATCATGT	ACAGGAAATT	CACGACGAAA	AGCGACGTCT	GGAGCCTGGG	GGTCTGTGTT	2580
TGGGAGATTT	TCACCTATGG	CAAACAGCCC	TGGTACCAGC	TGTCAAACAA	TGAGGTGATA	2640
GAGTGTATCA	CTCAGGGCCG	AGTCCTGCAG	CGACCCCGCA	CGTGCCCCCA	GGAGGTGTAT	2700
GAGCTGATGC	TGGGGTGTCT	GCAGCGAGAG	CCCCACATGA	GGAAGAACAT	CAAGGGCATC	2760
CATACCTTCC	TTCAGAACTT	GGCCAAAGCA	TCTCCGGTCT	ACCTGGACAT	TCTAGGCTAG	2820
GGCCCTTTTC	CCCAGACCGA	TCCTTCCCAA	CGTACTCCTC	AGACGGGCTG	AGAGGATGAA	2880
CATCTTTTAA	CTGCCGCTGG	AGGCCACCAA	GCTGCTCTCC	TTCACTCTGA	CAGTATTAAC	2940
ATCAAAGACT	CCGAGAAGCT	CTCGAGGGAA	GCAGTGTGTA	CTTCTTCATC	CATAGACACA	3000
GTATTGACTT	TTTTTTGGCA	TTATCTCTTT	CTCTCTTTCC	ATCTCCCTTG	GTTGTTCTTT	3060
TTTCTTTTTC	TAAATTTTCT	TTTTTCTCTT	TTTTTCTGTC	TTCCCTGCTT	CACGATTCTT	3120
ACCTTTTCTT	TTGAATCAAT	CTGGCTTCTG	CATTACTATT	AACCTCTGCAT	AGACAAAGGC	3180
CTTAACAAAC	GTAATTTGTT	ATATCAGCAG	ACACTCCAGT	TTGCCCACCA	CAACTAACAA	3240
TGCCTTGTTG	TATTCCTGCC	TTTGATGTGG	ATGAAAAAAA	GGGAAAACAA	ATATTTCACT	3300
TAAACTTTGT	CACTTCTGCT	GTACAGATAT	CGAGAGTTTC	TATGGATTCA	CTTCTATTTA	3360
TTTATTATTA	TTACTGTTCT	TATTGTTTTT	GGATGGCTTA	AGCCTGTGTA	TAAAAAAGAA	3420
AACTTGTGTT	CAATCTGTGA	AGCCTTTATC	TATGGGAGAT	TAAACCAGAA	GAGAAAGAAG	3480
ATTTATTATG	AACCGGAATA	TGGGAGGAAC	AAAGACAACC	ACTGGGATCA	GCTGGTGTCA	3540
GTCCCTACTT	AGGAAATACT	CAGCAACTGT	TAGCTGGGAA	GAATGTATTC	GGCACCTTCC	3600
CCTGAGGACC	TTTCTGAGGA	GTAAAAAGAC	TACTGGCCTC	TGTGCCATGG	ATGATTCTTT	3660
TCCCATCACC	AGAAATGATA	GCGTGCAGTA	GAGAGCAAAG	ATGGCTT		

Seq ID NO: 595 Protein sequence  
Protein Accession #: NP\_006171.1

75  
80  
85

1	11	21	31	41	51	
MSSWIRWHGP	AMARLWGFCW	LVVGFWRAAF	ACPTSCKCSA	SRIWCSDPSP	GIVAFPRLEP	60
NSVDPENITE	IFIANQKRLE	IINEDDVEAY	VGLRNLITVD	SGLKFAVAKA	FLKNSNLQHI	120
NFTRNKLTSL	SRKHFRHLDL	SELILVGNPF	TCSCDIMWIK	TLQEAKESSPD	TQDLYCLNES	180
SKNIPLANLQ	IPNCGLP SAN	LAAPNLVVEE	GKSITLSCSV	AGDPVPNMYW	DVGNLVSKHM	240
NETSHTQGS	LITNISSDDS	GKQISCVAEN	LVGEDQDSVN	LTVHFAPTIT	FLESPTSDDH	300
WCIPFTVKGN	PKPALQWFYN	GAILNESKYI	CTKIHVNTHT	EYHGCLQLDN	PTHMNGDYT	360
LIKNEYGKD	EKQISAHFMG	WPGIDDGANP	NYPDVIYEDY	GTAANDIGDT	TNRSNEIPST	420
DVTDKTGREH	LSVYAVVVIA	SVVGFCLLVM	LFLLLKLARH	KFGMKGPASV	ISNDDDSASP	480
LHHISNGSNT	PSSSEGGPDA	VIIGMTKIPV	IENPQYFGIT	NSQLKPDFTV	QHIKRHNIVL	540
KRELGEAGFG	KVFLAECYNL	CPEQDKILVA	VKTLKDASDN	ARKDFHREAE	LLTNLQHEHI	600
VKFYGVCEVE	DPLIMVFEYM	KHGDNLKFLR	AHGPDAVLMA	EGNPPTLTQ	SQMLHIAQQI	660
AAGMVYLASQ	HFVHRDLATR	NCLVGENLLV	KIGDFGMSRD	VYSTDYRVG	GHTMLPIRWM	720



PPESIMYRKF TTESDVWSLG VVLWEIFTYG KQPWYQLSNN EVIECITQGR VLQRPRTCPQ 780  
EVYELMLGCW QREPHMRKNI KGIHTLLQNL AKASPVYLDI LG

Seq ID NO: 596 DNA sequence  
Nucleic Acid Accession #: AF410899  
Coding sequence: 483..2999

10 1 11 21 31 41 51  
| | | | | |  
GGGAGCAGGA GCCTCGCTGG CTGCTTCGCT CGCGCTCTAC GCGCTCAGTC CCCGGCGGTA 60  
GCAGGAGCCT GGACCCAGGC GCCGGCGGCG GCGGTGAGGC GCCGGAGCCC GGCCTCGAGG 120  
TGCATACCGG ACCCCCATTC GCATCTAACA AGGAATCTGC GCCCCAGAGA GTCCCGGACG 180  
CCGCCGGTCT GTGCCCGGCG CGCCGGGCCA TGCAGCGACG GCCGCCGCGG AGCTCCGAGC 240  
AGCGGTAGCG CCCCCCTGTA AAGCGGTTCT CTATGCCGGG ACCACTGTGA ACCCTGCCGC 300  
15 CTGCCGGAAC ACTCTTCGCT CCGGACCAGC TCAGCCTCTG ATAAGCTGGA CTCGGCACGC 360  
CCGCAACAAG CACCGAGGAG TTAAGAGAGC CGCAAGCGCA GGAAGGCCT CCCCGCACGG 420  
GTGGGGGAAA GCGGCCGGTG CAGCGCGGGG ACAGGCACTC GGGCTGGCAC TGGCTGCTAG 480  
GGATGTCGTC CTGGATAAGG TGGCATGGAC CCGCCATGGC GCGGCTCTGG GGCTTCTGCT 540  
GGCTGGTTGT GGGCTTCTGG AGGGCCGCTT TCGCCTGTCC CACGTCCTGC AAATGCAGTG 600  
20 CCTCTCGGAT CTGGTGCAGC GACCCCTCTC CTGGCATCGT GGCATTTCCT AGATTGGAGC 660  
CTAACAGTGT AGATCCTGAG AACATCACCC ATATTTTCAT CGCAAAACCAG AAAAGGTTAG 720  
AAATCATCAA CTAAGATGAT GTTGAAGCTT ATGTGGGACT GAGAAATCTG ACAATTGTGG 780  
ATTCTGGATT AAAATTTGTG GCTCATAAAG CATTTCTGAA AAACAGCAAC CTGCAGACA 840  
TCAATTTTAC CCGAAACAAA CTGACGAGTT TGTCTAGGAA ACATTTCCGT CACCTTGACT 900  
25 TGTCTGAACT GATCCTGGTG GGCAATCCAT TTACATGCTC CTGTGACATT ATGTGGATCA 960  
AGACTCTCCA AGAGGCTAAA TCCAGTCCAG ACACTCAGGA TTTGTACTGC CTGAATGAAA 1020  
GCAGCAAGAA TATTCCTCTG GCAAACCTGC AGATACCCAA TTGTGGTTTG CCATCTGCAA 1080  
ATCTGGCCGC ACCTAACCTC ACTGTGGAGG AAGGAAAGTC TATCACATTA TCCTGTAGTG 1140  
TGGCAGGTGA TCCGGTTCCT AATATGTATT GGGATGTTGG TAACCTGGTT TCCAAACATA 1200  
30 TGAATGAAAC AAGCCACACA CAGGGCTCCT TAAGGATAAC TAACATTTCA TCCGATGACA 1260  
GTGGGAAGCA GATCTCTTGT GTGGCGGAAA ATCTTGTAAG AGAAGATCAA GATTCTGTCA 1320  
ACCTCACTGT GCATTTTGCA CCAACTATCA CATTTCTCGA ATCTCCAACC TCAGACCACC 1380  
ACTGGTGCAT TCCATTCACT GTGAAAGGCA ACCCCAAACC AGCGCTTCAG TGGTTCTATA 1440  
ACGGGGCAAT ATTGAATGAG TCCAAATACA TCTGTACTAA AATACATGTT ACCAATCACA 1500  
35 CGGAGTACCA CGGCTGCCCT CAGCTGGATA ATCCCACTCA CATGAACAAT GGGGACTACA 1560  
CTCTAATAGC CAAGAATGAG TATGGGAAGG ATGAGAAACA GATTTCTGCT CACTTCATGG 1620  
GCTGGCCTGG AATTGACGAT GGTGCAAACC CAAATTATCC TGATGTAATT TATGAAGATT 1680  
ATGGAACCTG AGCGAATGAC ATCGGGGACA CCACGAACAG AAGTAATGAA ATCCCTTCCA 1740  
CAGACGTCAC TGATAAAACC GGTCCGGGAA ATCTCTCGGT CTATGCTGTG GTGGTGATTG 1800  
40 CGTCTGTGGT GGGATTTTGG CTTTGGTAA TCTTGTCTT GCTTAAGTTG GCAAGACACT 1860  
CCAAGTTTGG CATGAAAGAT TTCTCATGGT TTGGATTTGG GAAAGTAAAA TCAAGACAAG 1920  
GTGTTGGCCC AGCCTCCGTT ATCAGCAATG ATGATGACTC TGCCAGCCCA CTCCATCACA 1980  
TCTCCAATGG GAGTAACACT CCATCTTCTT CGGAAGGTGG CCCAGATGCT GTCATTATTG 2040  
GAATGACCAA GATCCCTGTC ATTGAAATC CCCAGTACTT TGGCATCACC AACAGTCAGC 2100  
45 TCAAGCCAGA CACATTTGTT CAGCACATCA AGCGACATAA CATTTGTTCTG AAAAGGGAGC 2160  
TAGGCGAAGG AGCCTTTGGA AAAGTGTTC CAGTGAATG CTATAACCTC TGTCTGAGC 2220  
AGGACAAGAT CTTGGTGGCA GTGAAGACCC TGAAGGATGC CAGTGACAAT GCACGCAAGG 2280  
ACTTCCACCG TGAGGCCGAG CTCCTGACCA ACCTCCAGCA TGAGCACATC GTCAAGTTCT 2340  
ATGGCGTCTG CGTGGAGGGC GACCCCTCA TCATGGTCTT TGAGTACATG AAGCATGGGG 2400  
50 ACCTCAACAA GTTCCTCAGG GCACACGGCC CTGATGCCGT GCTGATGGCT GAGGGCAACC 2460  
CGCCACGGA ACTGACGCAG TCGCAGATGC TGCATATAGC CCAGCAGATC GCCGCGGGCA 2520  
TGGTCTACCT GCGTCCAG CACTTCGTG ACCCGGATTG GGCCACCAGG AACTGCCCTGG 2580  
TCGGGGAGAA TTTGCTGGTG AAAATCCGGG ACTTTGGGAT GTCCCGGGAC GTGTACAGCA 2640  
CTGACTACTA CAGGGTCCGT GGCCACACAA TGCTGCCCAT TCGCTGGATG CCTCCAGAGA 2700  
55 GCATCATGTA CAGGAAATTC ACGACGAAA GCGACGCTG GAGCCTGGGG GTCGTGTTGT 2760  
GGGAGATTTT CACCTATGGC AAACAGCCCT GGTACCAGCT GTCAAACAAT GAGGTGATAG 2820  
AGTGATCAC TCAGGGCCGA GTCCCTGCAGC GACCCCGCAC GTGCCCCCAG GAGGTGTATG 2880  
AGCTGATGCT GGGGTGCTGG CAGCGAGAGC CCCCATGAG GAAGAACATC AAGGGCATCC 2940  
ATACCTCTCT TCAGAACTTG GCCAAGGAT CTTCCGTTCTA CCTGGACATT CTAGGCTAGG 3000  
60 GCCCTTTTCC CCAGACCGAT CCTTCCCAAC GTACTCTCA GACGGGCTGA GAGGATGAAC 3060  
ATCTTTTAAC TGCCGCTGGA GGCCACCAAG CTGCTCTCCT TCACTCTGAC AGTATTAACA 3120  
TCAAAGACTC CGAGAAGCTC TCGAGGGAAG CAGTGTGTAC TTCTTCATCC ATAGACACAG 3180  
TATTGACTTC TTTTGGCAT TATCTCTTTC TCTCTTTCCA TCTCCCTTGG TTGTTCTTTT 3240  
TTCTTTTCTT AAATTTCTT TTTCTTCTT TTTTTCGTCT TCCCTGCTTC ACGATTCTTA 3300  
65 CCCTTTCTTT TGAATCAAT TGGCTTCTG ATTACTATTA ACTCTGCATA GACAAAGGCC 3360  
TTAACAACCG TAATTTGTTA TATCAAGACA CACTCCAGTT TGCCACCAC AACTAACAAT 3420  
GCCTTGTTGT ATTCCTGCCT TTGATGTGGA TGAAAAAAG GGAACAACA TATTTCACTT 3480  
AACTTTGTC ACTTCTGCTG TACAGATATC GAGAGTTTCT ATGGATTAC TTCTATTTAT 3540  
70 TTATTATTAT TACTGTTCTT ATGTGTTTGG GATGGCTTAA GCCTGTGTAT AAAAAAGAAA 3600  
ACTTGTGTTT AATCTGTGAA GCCTTTATCT ATGGGAGATT AAAACCAGAG AGAAAGAAGA 3660  
TTTATTATGA ACCGCAATAT GGGAGGAACA AAGACAACCA CTGGGATCAG CTGGTGTGAG 3720  
TCCCTACTTA GGAAATACTC AGCAACTGTT AGCTGGGAAG AATGTATTCG GCACCTTCCC 3780  
CTGAGGACCT TTCTGAGGAG TAAAAAGACT ACTGGCCTCT GTGCCATGGA TGATTCTTTT 3840  
CCCATCACCA GAAATGATAG CGTGCACTAG AGAGCAAAGA TGGCTTCCGT GAGACACAAG 3900  
75 ATGGCGCATA GTGTGCTCGG ACACAGTTT GTCTTCGTAG GTTGTGATGA TAGCACTGGT 3960  
TTGTTTCTCA AGCGCTATCC ACAGAACCTT TGTCAACTTC AGTTGAAAAG AGGTGGATTTC 4020  
ATGTCCAGAG CTCATTTCCG GGTGAGGTGG GAAAGCC

Seq ID NO: 597 Protein sequence  
Protein Accession #: AAL67965.1

80 1 11 21 31 41 51  
| | | | | |  
MSSWIRWHGP AMARLWGFCW LVVGFWRAAF ACPTSCCKSA SRIWCSDPSP GIVAFPRLEP 60  
85 NSVDPENITE IFIANQKRLE IINEDDVEAY VGLRNLITVD SGLKFVAHKA FLKNSNLQHI 120  
NFRNKLTSI SRKHFRHLDL SELILVGNPF TCSCDIMWIK TLQEAQSSPD TQDLVCLNES 180  
SKNIPLANLQ IPNCGLPSAN LAAPNLTVEE GKSITLSCSV AGDPVPNMYW DVGNLVSXHM 240



NETSHTQSSL RITNISSDDS GKQISCVAEN LVGEDQDSVN LTVHFAPTIT FLESPTSDDH 300  
WCIPFTVKGN PKPALQWFYN GAILNESKYI CTKIHVTNHT EYHGCLQLDN PTHMNGDYT 360  
LIAKNEYGKD EKQISAHFMG WPGIDDGANP NYPDVIYEDY GTAANDIGDT TNRSNEIPST 420  
5 DVTDKTGREH LSVYAVVIA SVVGFCLLVM LFLKLARHS KFGMKDFSWF GFGKVKSRQG 480  
VGPASVISND DDSASPLHHI SNGSNTSPSSS EGGPDAVIIG MTKIPVIENP QYFGITNSQL 540  
KPDFTFVQHIK RHNIVLKREL GEGAFGKVFL AECYNLCPEQ DKILVAVKTL KDASDNARKD 600  
FHREAELLTN LQEHIVKFY GVCVEGDPLI MVFEYMKHGD LNKFLRAHGP DAVLMAEGNP 660  
PTELTQSQML HIAQQIAAGM VYLASQHFVH RDLATRNCLV GENLLVKIGD FGMSRDVYST 720  
10 DYYRVGGHTM LPIRWMPPE S IMYRKFTTES DVWSLGVVLW EIFTYQKQPW YQLSNNEVIE 780  
CITQGRVLQR PRTCPQEVYE LMLGCWQREP HMRKNIKGIH TLLQNLAKAS PVYLDILG

Seq ID NO: 598 DNA sequence  
Nucleic Acid Accession #: AB052906  
Coding sequence: 74..814

1 11 21 31 41 51  
| | | | |  
AAAACCTTGA GGTGATTTCAT CTTCAGGCT CTCTTCCAT CAAGTCTCTC CTCCCTAGCG 60  
20 CTCTGGGTCC TTAATGGCAG CAGCCGCCGC TACCAAGATC CTTCTGTGCC TCCCGCTTCT 120  
GCTCCTGCTG TCCGGCTGGT CCCGGGCTGG GCGAGCCGAC CCTCACTCTC TTTGCTATGA 180  
CATCACCGTC ATCCCTAAGT TCAGACCTGG ACCACGGTGG TGTGCGGTTT AAGGCCAGGT 240  
GGATGAAAAG ACTTTTCTTC ACTATGACTG TGGCAACAAG ACAGTCACAC CTGTCACTCC 300  
CCTGGGGAAG AAATAAATG TCACAACGGC CTGGAAAGCA CAGAACCAG TACTGAGAGA 360  
25 GGTGGTGGAC ATACTTACAG AGCAACTGCG TGACATTGAG CTGGAGAATT ACACACCCAA 420  
GGAACCCCTC ACCCTGCAGG CCAGGATGTC TTTGTGAGCAG AAAGCTGAAG GACACAGCAG 480  
TGGATCTTGG CAGTTTCAGT TCGATGGGCA GATCTTCTCT CTCTTTGACT CAGAGAAGAG 540  
AATGTGGACA ACGGTTTCATC CTGGAGCCAG AAAGATGAAA GAAAAGTGGG AGAATGACAA 600  
GGTTGTGGCC ATGTCTTCTC ATTACTTCTC AATGGGAGAC TGTATAGGAT GGCTTGAGGA 660  
30 CTTCTTGATG GGCATGGACA GCACCTTGGA GCCAAGTGCA GGAGCACCAC TCGCCATGTC 720  
CTCAGGCACA ACCCAACTCA GGGCCACAGC CACCACCCTC ATCCTTTGCT GCCTCCTCAT 780  
CATCCTCCCC TGCTTCATCC TCCCTGGCAT CTGAGGAGAG TCCTTTAGAG TGACAGGTTA 840  
AAGCTGATAC CAAAAGGCTC CTGTGAGCAC GGTCTTGATC AAACCTCGCC TTCTGTCTGG 900  
CCAGCTGCCC AGGACCTACG GTGTATGTCC AGTGGCCTCC AGCAGATCAT GATGACATCA 960  
35 TGGACCCAAT AGCTCATTC CAGCTTGTAT TCCTTTTGCC AACAATTTTA CCAGCAGTTA 1020  
TACCTAACAT ATTATGCAAT TTTCTCTTGG TGCTACCTGA TGAATTCCT GCACTTAAAG 1080  
TTCTGGCTGA CTAAACAAGA TATATCATTT TCTTTCTTCT CTTTTTGTTC GGAAATCAA 1140  
GTACTTCTTT GAATGATGAT CTCTTCTTGG CAAATGATAT TGTCAGTAAA ATAATCACGT 1200  
TAGACTTCAG ACCTCTGGGG ATTCTTTCCG TGTCCTGAAA GAGAATTTT AAATTATTTA 1260  
40 ATAAGAAAAA ATTTATATTA ATGATTGTTT CCTTTAGTAA TTTATTGTTT TGTACTGATA 1320  
TTTAAATAAA GAGTTCTATT TCCCAAAAAA AAAAAAAAAA AA

Seq ID NO: 599 Protein sequence  
Protein Accession #: BAB61048.1

1 11 21 31 41 51  
| | | | |  
MAAAAATKIL LCLPLLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGVVDEKT 60  
FLHYDCGKNT VTPVSPGK LNVTTAWKAQ NPVLREVVDI LTEQLRDIQL ENYTPKEPLT 120  
50 LQARMSCEQK AEGHSSGSWQ FSPDQIPLL FDSEKRMWTT VHPGARKMKE KWENDKVVM 180  
SFHYFSMGDC IGWLEDFLMG MDSTLEPSAG APLAMSSGTT QLRATATTLI LCCLLIILPC 240  
FILPGI

Seq ID NO: 600 DNA sequence  
Nucleic Acid Accession #: NM\_001898.1  
Coding sequence: 57..482

1 11 21 31 41 51  
| | | | |  
GGCTCTCACC CTCCTCTCCT GCAGCTCCAG CTTTGTGCTC TGCCTCTGAG GAGACCATGG 60  
60 CCCAGTATCT GAGTACCCTG CTGCTCCTGC TGGCCACCCT AGCTGTGGCC CTGGCCTGGA 120  
GCCCCAAGGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCAATGATG 180  
AGTGGGTACA GCGTGCCCTT CACTTCGCCA TCAGCGAGTA TAACAAGGCC ACCAAAGATG 240  
ACTACTACAG ACGTCCGCTG CCGGTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGGTGA 300  
ATTACTTCTT CGACGTAGAG GTGGGCCGCA CCATATGTAC CAAGTCCCAG CCCAACTTGG 360  
65 ACACCTGTGC CTTCCATGAA CAGCCAGAAC TGCAGAAGAA ACAGTTGTGC TCTTTCGAGA 420  
TCTACGAAGT TCCCTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCCT 480  
AGGGATCTGT GCCAGGCCAT TCGCACCAGC CACCACCCAC TCCCACCCCT TGTAGTGCTC 540  
CCACCCCTGG ACTGGTGGCC CCCACCCCTG GGGAGGCCTC CCCATGTGCC TCGCCCAAGA 600  
70 GACAGACAGA GAAGGCTGCA GGAGTCCCTT GTTGCTCAGC AGGGCGCTCT GCCCTCCCTC 660  
CTTCTTCTT GCTTCTAATA GCCCTGGTAC ATGGTACACA CCCCCCACC TCCTGCAATT 720  
AAACAGTAGC ATCGCC

Seq ID NO: 601 Protein sequence  
Protein Accession #: NP\_001889.1

1 11 21 31 41 51  
| | | | |  
MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWVQALHF AISEYNKATK 60  
80 DDYYRRPLRV LRARQTVGG VNYFFDVEVG RTICTKSQPN LDTCAFHEQP ELQKKQLCSF 120  
EIYEVWENR RSLVKSRCQE S

Seq ID NO: 602 DNA sequence  
Nucleic Acid Accession #: NM\_003976.2  
Coding sequence: 299..961

1 11 21 31 41 51  
| | | | |



CTCTGAGCTT CTCTGAGCCT TGTTCGCTCA TCTGGAAAAA GGGGATTAAA CCATTACCT 60  
CATGGAGTTG TGAAAGAATA GCTGCAAAGC ACCTAACACA TAGTAAGGTT CCCAGTGCAG 120  
CTACTTCTGC TGGGTTGAGT CTAGCTGTGT AGGCCCCCTG TTCCTCACCT GGAGAACTG 180  
GGGTGGCAGG CCGGTCCCCC ACAAAGATA ACTCATCTCT TAATTGCAA GCTGCCTCAA 240  
CAGGAGGGTG GGGGAACAGC TCAACAATGG CTGATGGGCG CTCCTGGTGT TGATAGAGAT 300  
GGAACCTTGA CTTGGAGGCC TCTCCACGCT GTCCCACTGC CCCTGGCCTA GGCGGCAGCC 360  
TGCCCTGTGG CCCACCCTGG CCGCTCTGGC TCTGCTGAGC AGCGTCGCAG AGGCCTCCCT 420  
GGGCTCCGCG CCGCGCAGCC CTGCCCCCGG CGAAGGCCCC CCGCCTGTCC TGGCGTCCCC 480  
CGCCGGCCAC CTGCGGGGGG GACGCACGGC CCGCTGGTGC AGTGGAAGAG CCCGCGGCC 540  
GCCGCGCAG CTTCTCGGC CCGCGCCCCC GCCGCTGCA CCCCCATCTG CTCTCCCCG 600  
CGGGGGCCGC GCGGCGCGGG CTGGGGGGCC GGGCAGCCGC GCTCGGGCAG CGGGGGCGCG 660  
GGGCTGCCGC CTGCGCTCGC AGCTGGTGCC GGTGCGCGCG CTCGGCCTGG GCCACCGCTC 720  
CGACGAGCTG GTGCGTTTCC GCTTCTGCAG CCGCTCCTGC CGCCGCGCGC GCTCTCCACA 780  
CGACCTCAGC CTGGCCAGCC TACTGGGCGC CGGGGCCCTG CGACCGCCCC CGGGCTCCCG 840  
GCCCGTCAGC CAGCCCTGCT GCCGACCCAC GCGCTACGAA GCGGTCTCCT TCATGGACGT 900  
CAACAGCACC TGGAGAACCG TGGACCGCCT CTCCGCCACC GCCTGCGGCT GCCTGGGCTG 960  
AGGGCTCGCT CCAGGGCTTT GCAGACTGGA CCCTTACCGG TGGCTCTTCC TGCTGGGAC 1020  
CCTCCCGCAG AGTCCCACTA GCCAGCGGCC TCAGCCAGGG ACGAAGGCCT CAAAGCTGAG 1080  
AGGCCCCAC CCGTGGGTGA TGGATATCAT CCCCAGACAG GTGAAGGGAC AACTGACTAG 1140  
CAGCCCCAGA GCCCTCACCC TGCGGATCCC AGCCTAAAAG ACACCAGAGA CCTCAGCTAT 1200  
GGAGCCCTTC GGACCCACTT CTCACAGACT CTGGCACTGG CCAGGCCTCG AACCTGGGAC 1260  
CCCTCCTCTG ATGAACACTA CAGTGGCTGA GGCATCAGCC CCCGCCAGG CCCTGTAGGG 1320  
ACAGCATTTG AAGGACACAT ATTGCAGTTG CTTGGTTGAA AGTGCCTGTG CTGGAAGTGG 1380  
CCTGTACTCA CTCATGGGAG CTGGCCCC

Seq ID NO: 603 Protein sequence  
Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
| | | | |  
MELGLGGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120  
RGCLRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180  
RPVSPQCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 604 DNA sequence  
Nucleic Acid Accession #: NM\_057091.1  
Coding sequence: 783..1445

1 11 21 31 41 51  
| | | | |  
ACTGGCCGCT GAGAGAAGAA TCGGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60  
GGACCCCCAA ATCTGCACGT ACCAGCAGTC AGCCGCCCCA CGCAGGGACC GGCTTACCCC 120  
TCGCTCCCCG CCCTCACTCA CTTTCTCCCG CCTCGGCCC GGCCTCCAG CTCTCTACTT 180  
CGCGTGTCTA CAAACTCAAC TCCCGGTTTC CGTGCCTCTC CACCGCTCGA GTTCTCTACT 240  
CTCCATATCC GAGGGGCCCC TCCAGCATC TACCCCTCTC CCAACCTCGG GGGACCTAGC 300  
CAAGCTAGGG GGGACTGGAT CCGACGGGTG GAGCAGCCAG GTGAGCCCCG AAAGGTGGGG 360  
CGGGGCAGGG GCGCTCCCAG CCCCACCCCG GGATCTGGTG ACGCTGGGGC TGGAAATTTGA 420  
CACCGGACGG CTGCGGCGGC GGGCAGGAGG CTGCTGAGGG ATGGAGTTGG GCCCGCCCC 480  
CAGACAAGGC CCGGGGGCTC CGCCAGCAGC AGGTCCCTCG GGGCCAGCC CTCGCTGCCA 540  
CCCGGGCCTG GAGCCCCACA CCGAGGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600  
TAAAGAGGC ACTGCCAGGT GTACAGTCTT GGCATGCGC TGTTCGAGCT TCGGGGGAGA 660  
GCCAGCACT GGTCCCCGGA AAGTGCCTA GAAGAACAAG GTGCAGGACC CCGTCTGCC 720  
TCAACAGGAG GGTGGGGGAA CAGCTCAACA ATGGCTGATG GCGCTCCTG GTGTTGATAG 780  
AGATGGAAGT TGGACTTGGA GGCTCTCCA CGCTGTCCCA CTGCCCCTGG CCTAGGCGGC 840  
AGCCTGCCCT GTGGCCACC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCCT 900  
CCCTGGGCTC CGCGCCCCGC AGCCCTGCCC CCGCGAAGG CCCCCGCTT GTCCTGGCGT 960  
CCCCGCGCG CCACCTGCCG GGGGACGCA CCGCCGCTG GTGCAGTGA AGAGCCCGC 1020  
GGCCGCGCG CAGCCCTTCT CGGCCCGCC CCGCCGCGC TGCACCCCA CTGCTCTTC 1080  
CCCGCGGGG CCGCGCGCG CCGGCTGGGG GCCCGGGCAG CCGCGCTCGG GCAGCGGGG 1140  
CGCGGGGCTG CCGCTGCGC TCGCAGCTGG TGCCGGTGCG CCGCTCGGC CTGGGCCACC 1200  
GCTCCGACGA GCTGGTGCGT TTCCGCTTCT GCAGCGGCTC CTGCGCGCGC GCGCGCTCTC 1260  
CACACGACCT CAGCCTGGCC AGCCTACTGG GCGCCGGGGC CCTGCGACCG CCCCCGGGCT 1320  
CCCGGCCGT CAGCCAGCCC TGCTGCCGTA CCAAGCGGTC CGAAGCGGTC TCCTTCATGG 1380  
ACGTCAACAG CACCTGGAGA ACCCTGGACC GCCTCTCCG CACCGCTGC GCTGCTGG 1440  
GCTGAGGGCT CGCTCCAGGG CTTTGCAGAC TGGACCTTA CCGGTGGCTC TTCCTGCTG 1500  
GGACCTCCC GCAGAGTCCC ACTAGCCAGC GGCTCAGCC AGGGACGAAG GCCTCAAAGC 1560  
TGAGAGGCCC CTACCGGTGG GTGATGGATA TCATCCCCGA ACAGGTGAAG GGACAACTGA 1620  
CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG 1680  
CTATGGAGCC CTTGCGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG 1740  
GGACCCCTCC TCTGATGAAC ACTACAGTGG CTAGGCAATC AGCCCCCGC CAGGCCCTGT 1800  
AGGGACAGCA TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAAGTGCC TGTGCTGGAA 1860  
CTGGCCTGTA CTCACTCATG GGAGCTGGCC CC

Seq ID NO: 605 Protein sequence  
Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
| | | | |  
MELGLGGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120  
RGCLRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180  
RPVSPQCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 606 DNA sequence  
Nucleic Acid Accession #: NM\_057160.1



Coding sequence: 1..714

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      ATGCCCGGCC TGATCTCAGC CCGAGGACAG CCCCTCCTTG AGGTCCTTCC TCCCCAAGCC 60
      CACCTGGGTG CCCTCTTCT CCCTGAGGCT CCACTTGGTC TCTCCGCGCA GCCTGCCCTG 120
      TGGCCACCC TGGCCGCTCT GGCTCTGCTG AGCAGCGTCG CAGAGGCCTC CCTGGGCTCC 180
      GCGCCCCGCA GCCCTGCCCC CCGCGAAGGC CCCCCGCTG TCCTGGCGTC CCCC GCCGGC 240
      CACCTGCCGG GGGGACGCAC GGCCCGCTGG TGCACTGGAA GAGCCCGGCG GCCGCCGCCG 300
10     CAGCCTTCTC GGCCCGCGCC CCCGCCGCTG GCACCCCAT CTGCTCTTCC CCGCGGGGGC 360
      CGCGCGGCGC GGGCTGGGGG CCCGGGCGAG CGCGCTCGGG CAGCGGGGGC GCGGGGCTGC 420
      CGCCTGCGCT CGCAGCTGGT GCCGGTGGCG GCGCTCGGCC TGGGCCACCG CTCCGACGAG 480
      CTGGTGCCTT TCCGCTTCTG CAGCGGCTCC TGCCGCGCGG CGCGCTCTCC ACACGACCTC 540
      AGCCTGGCCA GCCTACTGGG CGCCGGGGCC CTGCGACCGC CCCC GGCTC CCGGCCCGTC 600
15     AGCCAGCCCT GCTGCCGACC CACGCGCTAC GAAGCGGTCT CCTTCATGGA CGTCAACAGC 660
      ACCTGGAGAA CCGTGGACCG CCTCTCCGCC ACCGCCTGCG GCTGCCTGGG CTGAGGGCTC 720
      GCTCCAGGGC TTTGCAGACT GGACCTTAC CCGTGGCTCT TCCTGCCTGG GACCCTCCCG 780
      CAGAGTCCCA CTAGCCAGCG GCCTCAGCCA GGGACGAAGG CCTCAAAGCT GAGAGGCCCC 840
      TACCGGTGGG TGATGGATAT CATCCCCGAA CAGGTGAAGG GACAACTGAC TAGCAGCCCC 900
20     AGAGCCCTCA CCCTGCGGAT CCCAGCCTAA AAGACACAG AGACCTCAGC TATGGAGCCC 960
      TTCGGACCCA CTTCTCACAG ACTCTGGCAG TGGCCAGGCC TCGAACCTGG GACCCCTCCT 1020
      CTGATGAACA CTACAGTGGC TGAGGCATCA GCCCCCGCCC AGGCCCTGTA GGGACAGCAT 1080
      TTGAAGGACA CATATTGCAG TTGCTTGGTT GAAAGTGCTT GTGCTGGAAC TGGCCTGTAC 1140
      TCACTCATGG GAGCTGGCCC C

```

Seq ID NO: 607 Protein sequence  
Protein Accession #: NP\_476501.1

```

30     1      11      21      31      41      51
      |      |      |      |      |      |
      MPGLISARGQ PLLEVLPPQA HLGALFLPEA PLGLSAQPAL WPTLAALALL SSVAEASLGS 60
      APRSPAPREG PPPVLASPAG HLPGGRTARW CSGRRARRPP QPSRPAPPPP APPSALPRGG 120
      RAARAGGPGS RARAAGARGC RLRSQLVFVR ALGLGHRSD LVRFRFCSGS CRRARSPHDL 180
35     SLASLLGAGA LRPPPGSRPV SQPCCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG

```

Seq ID NO: 608 DNA sequence  
Nucleic Acid Accession #: NM\_057090.1  
Coding sequence: 29..715

```

40     1      11      21      31      41      51
      |      |      |      |      |      |
      CTGATGGGCG CTCCTGGTGT TGATAGAGAT GGAACCTGGA CTTGGAGGCC TCTCCACGCT 60
      GTCCCACTGC CCCTGGCCTA GGCGGCAGGC TCCACTTGGT CTCTCCGCGC AGCCTGCCCT 120
      GTGGCCACCC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCCT CCCTGGGCTC 180
45     CGCGCCCCGC AGCCCTGCCC CCCGCGAAGC CCCCCGCTG GTCCTGGCGT CCCC GCCCGC 240
      CCACCTGCCG GGGGGACGCA CGGCCCGCTG GTGCAGTGGA AGAGCCCGGC GGCCGCCGCC 300
      GCAGCCTTCT CGGCCCGCGC CCCC GCCGCC TGACCCCCCA TCTGCTCTTC CCCGCCGGGG 360
      CCGCGCGGCG CGGGCTGGGG GCGCGGCGAG CCGCGCTCGG GCAGCGGGGG CGCGGGGCTG 420
      CCGCCTGCGC TCGCAGCTGG TGCCGCTGCG CGCGCTCGGC CTGGGCCACC GCTCCGACGA 480
50     GCTGGTGCCT TCCCGCTTCT GCAGCGGCTC CTGCCGCCGC GCGCGCTCTC CACACGACCT 540
      CAGCCTGGCC AGCCTACTGG GCGCCGGGGC CCTGCGACCG CCCC GGCTC CCGGCCCGCT 600
      CAGCCAGCCC TGCTGCCGAC CCACGCGTCA CGAAGCGGTC TCCTTCATGG ACGTCAACAG 660
      CACCTGGAGA ACCGTGGACC GCCTCTCCGC CACCGCCTGC GGCTGCCTGG GCTGAGGGCT 720
      CGCTCCAGGG CTTTGCAGAC TGGACCCTTA CCGGTGGCTC TTCCTGCCTG GGACCCCTCC 780
55     GCAGAGTCCC ACTAGCCAGC GGCTCAGCC AGGGACGAAG GCCTCAAAGC TGAGAGGCCC 840
      CTACCGGTGG GTGATGGATA TCATCCCCGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900
      CAGAGCCCTC ACCCTGCGGA TCCCAGCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960
      CTTCCGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG GGACCCCTCC 1020
      TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT AGGGACAGCA 1080
60     TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAAGTGCC TGTGCTGGAA CTGGCCTGTA 1140
      CTCCTCATG GGAGCTGGCC CC

```

Seq ID NO: 609 Protein sequence  
Protein Accession #: NP\_476431.1

```

65     1      11      21      31      41      51
      |      |      |      |      |      |
      MELGLGGLST LSHCPWPRRQ APLGLSAQPA LWPTLAALAL LSSVAEASLG SAPRSPAPRE 60
      GPPPVLASPA GHLPGGRTAR WSGRRARRPP PQPSRPAPPP PAPPALPRG GRAARAGGPG 120
70     SRARAAGARG CRLRSQLVFV RALGLGHRSD ELVRFRFCSG SCRRARSPHD LSLASLLGAG 180
      ALRPPPGSRP VSQPCCRPTR YEAVSFMDVN STWRTVDRLS ATACGCLG

```

Seq ID NO: 610 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1746

```

75     1      11      21      31      41      51
      |      |      |      |      |      |
      ATGCCACTGA AGCATTATCT CCTTTTGCTG GTGGGCTGCC AAGCCTGGGG TGCAGGGTTG 60
      GCCTACCATG GCTGCCCTAG CGAGTGTACC TGCTCCAGGG CCTCCAGGT GGAGTGCACC 120
      GGGGCACGCA TTGTGGCGGT GCCCACCCTT CTGCCCTGGA ACGCCATGAG CCTGCAGATC 180
      CTCAACACGC ACATCACTGA ACTCAATGAG TCCCCGTTCC TCAATATCTC AGCCCTCATC 240
80     GCCCTGAGGA TTGAGAAGAA TGAGCTGTCT CGCATCACGC CTGGGGCCTT CCGAAACCTG 300
      GGCTCGCTGC GCTATCTCAG CCTCGCCAAC AACAAGCTGC AGGTTCTGCC CATCGGCCTC 360
      TTCCAGGGCC TGGACAGCCT TGAGTCTCTC CTTCTGTCCA GTAACCAGCT GTTGAGATC 420
      CAGCCGGCCC ACTTCTCCCA GTGCAGCAAC CTCAAGGAGC TGCAGTTGCA CGGCAACCAC 480
85     CTGGAATACA TCCCTGACGG AGCCTTCGAC CACCTGGTAG GACTCACGAA GCTCAATCTG 540

```



	GGCAAGAATA	GCCTCACCCA	CATCTCACCC	AGGGTCTTCC	AGCACCTGGG	CAATCTCCAG	600
	GTCCTCCGGC	TGTATGAGAA	CAGGCTCACG	GATATCCCCA	TGGGCACTTT	TGATGGGCTT	660
	GTAAACCTGC	AGGAACCTGC	TCTACAGCAG	AACCAGATTG	GACTGCTCTC	CCCTGGTCTC	720
5	TTCCACAACA	ACCACAACCT	CCAGAGACTC	TACCTGTCCA	ACAACCACAT	CTCCAGCTG	780
	CCACCCAGCA	TCTTCATGCA	GCTGCCCCAG	CTCAACCGTC	TTACTCTCTT	TGGGAATTCC	840
	CTGAAGGAGC	TCTCTCTGGG	GATCTTCGGG	CCCATGCCCA	ACCTGCGGGA	GCTTTGGCTC	900
	TATGACAACC	ACATCTCTTC	TCTACCCGAC	AATGTCTTCA	GCAACCTCCG	CCAGTTGCAG	960
	GTCCCTGATT	TTAGCCGCAA	TCAGATCAGC	TTTATCTCCC	CGGGTGCCTT	CAACGGGCTA	1020
10	ACGGAGCTTC	GGGAGCTGTC	CCTCCACACC	AACGCACTGC	AGGACCTGGA	CGGGAATGTC	1080
	TTCCGCATGT	TGGCCAACCT	GCAGAACATC	TCCCTGCAGA	ACAATCGCCT	CAGACAGCTC	1140
	CCAGGGAATA	TCTTCGCCAA	CGTCAATGGC	CTCATGGCCA	TCCAGCTGCA	GAACAACCAG	1200
	CTGGAGAACT	TGCCCCCTCG	CATCTTCGAT	CACCTGGGGA	AACTGTGTGA	GCTGCGGCTG	1260
	TATGACAATC	CCTGGAGGTG	TGACTCAGAC	ATCCTTCCGC	TCCGCAACTG	GCTCCTGCTC	1320
	AACCAGCCTA	GGTTAGGGAC	GGACACTGTA	CCTGTGTGTT	TCAGCCGAGC	CAATGTCCGA	1380
15	GGCCAGTCCC	TCATTATCAT	CAATGTCAAT	GTTGCTGTTT	CAAGCGTCCA	TGTCCCTGAG	1440
	GTGCCTAGTT	ACCCAGAAAC	ACCATGGTAC	CCAGACACAC	CCAGTTACCC	TGACACCACA	1500
	TCCGTCTCTT	CTACCACTGA	GCTAACCAGC	CCTGTGGAAG	ACTACACTGA	TCTGACTACC	1560
	ATTGAGGTCA	CTGATGACCG	CAGCGTTTGG	GGCATGACCC	AGGCCGAGAG	CGGGCTGGCC	1620
20	ATTGCCGCCA	TTGTAATTGG	CATTGTGCGC	CTGGCCTGCT	CCCTGGCTGC	CTGCGTCGGC	1680
	TGTTGTCTGT	GCAAGAAGAG	GAGCCAAGCT	GTCCTGATGC	AGATGAAGGC	ACCCAATGAG	1740
	TGTTAAAGAG	GCAGGCTGGA	GCAGGGCTGG	GGAATGATGG	GACTGGAGGA	CCTGGGAATT	1800
	TCATCTTTCT	GCCTCCACCC	CTGGGTCCAT	GGAGCTTTCC	CGTGATTGCT	CTTTCTGGCC	1860
	CTAGATAAAG	GTGTGCCCTAC	CTCTTCCTGA	CTTGCCTGAT	TCTCCCGTAG	AGAAGCAGGT	1920
25	CGTGCCGGAC	CTTCCCTACAA	TCAGGAAGAT	AGATCCAAC	GGCCATGGCA	AAAGCCCTGG	1980
	GGATTTCCGA	TTCATACCCC	TGGGCTTCCT	TCGAGAGGGC	TCTTCTCCCA	AATCCTCCCC	2040
	ACCTGTCTCT	CAAGAACAGC	CTTCCCTGCG	CCCAGGCCCC	CTCCGGGCTC	CTGTAGACTC	2100
	AGTTAGTCCA	CAGCCTGCTC	ACTTCGTGGG	AATAGTTCTC	CGCTGAGATA	GCCCTCTCG	2160
	CCTAAGTATT	ATGTAAGTTG	ATTTCCTTTC	TTTTGTTTTCT	CTTGTGTTTG	CTATGGCTTG	2220
30	ACCCAGCATG	TCCCTCCTCAA	TGAAAGTTCT	CCCCTTGATT	TTCTGCTCCT	GAAGGCAGGG	2280
	TGAGTTCTCT	CCTCAAAGAA	GACTTCAAAC	CATTTAACTG	GTTTCTTAAG	AGCCGTCAAT	2340
	CAGCCTGGTT	TTGGGGATGC	TATGAAAGAG	AGAAGGAAAA	TCATGCCGCT	CAGTTCCCTGG	2400
	AGACAGAAGA	GCCGTCATCA	GTGTCTCACT	TGTGATTTTT	ATCTGGAAAA	GGAAGAAACA	2460
	CCCCAGCACA	GCAAGCTCAG	CCTTTTAGAG	AAGGATATTT	CCAAACTGCA	AACTTTGCTT	2520
35	TGAAAAGTTT	AGCCCTTTAA	GGAAATGAAAT	CATGTAGAAT	TTTGGACTTC	TAAAAACATT	2580
	AAAATCAGCT	TATTAATACG	GGATAGAGAA	AGAAATCTGG	TGCCCTGGGGG	TCCCTGTGTT	2640
	CACCCCTAGA	GTTTGTTTTA	AAATTTTAA	TTGAAGCATG	TGAAGTGATC	STGCAGAAAA	2700
	GTGGGAACAT	GATAGTGTAT	GGCTTGGTGG	ATTTTCACAA	ACTGAACATA	CCTGTGTAAT	2760
	CAGCATCTAG	ACCCAGACCC	AGAGCATCAC	AAATATCCCC	CATCCTGGGC	TTTTCCCGAG	2820
40	GGAGTGGGG	GCTTCTGAAG	ATGGACTTAC	CTGGGACCTG	CCCCCATGA	GCCAGGACGG	2880
	TCCCCCCACA	GTCAGCCTGT	GCAAGGCCCC	CGTGGCCAGG	GGTGGAGGAG	AATATGTGGG	2940
	TGTGGACAGG	ATGGGAGACT	GTGGCCTGAA	CAGGAGATTT	TATTATATCT	GGAGACCCTG	3000
	AGAGACCCTG	AGACCTGGGG	CACCATGGCT	GGCCAGGTCA	GAAGCATCCT	GACTGCAGAG	3060
	GTCCGTGCAG	CCACACCCTC	TTCCCTGCCA	GCAAGTTGTC	TGCGGCTCAT	CGGAGGCCCC	3120
45	TCCGCCTGGA	GCCTTCTATG	GACGTGATAT	GCCTGTATCT	GTTTTTAATT	TTCATTCTTC	3180
	ACTTAGGGGA	AGTGAATCG	CTCAGAGATG	AGATCCTTTA	ATTGAAAACG	AAGTGTAACG	3240
	GAATCTAGTG	GTCTTCTAAT	TGTTTAAAT	TCTCCATCAA	CATCACAGTC	AGCTGGCAGC	3300
	TGAACCTCAG	AATCTCACTT	ACAGCAGGCG	ACACGGGGGT	ACACCGATGG	GTCACACTGG	3360
	GTCTGGGGGC	TCCCTGGAGC	TCCTCCTGCG	TGTGGTCTGG	TTAGGAGTTG	AGTTGTTTGC	3420
50	TCCAGGGTTA	TTCTCCTCCT	CGAGTCACAG	TCACACGAAT	ACCTGCCCTC	TCTGGCTTTC	3480
	CTGCTATACA	CATATTCACA	TGGCGCTCAA	GAAGTTAGGC	TCATGGCAAC	GTGTGTCTTT	3540
	CTCTGGACAA	CTGGCCCACT	TTACAGTGAA	ATGGAGAATT	TCAGGTCTCC	ACGTCTGCCC	3600
	AGGAAAGAAC	TTCAGCTGAC	TCCACGGGGA	TCTGGAAATC	CACGACCAAT	CCCGATCGGC	3660
	TCTTATTAGC	TCCCGCTCC	ACAAGACACC	TGTGCTTTGG	AAATCCACCA	CCAATCCCGA	3720
55	TCGGCTCTTA	TTAGCTCCCC	GCTCCACAAG	ACACCTGTGA	TCTGGAAATC	TACCACCAAT	3780
	CCCGATCGGC	TCTTATTAGC	TCCCGCTCC	ACAAGACACC	TGTGACATCC	TCCAGGGCCA	3840
	CAGGAGCACG	TGCTGACCAG	TTTTCCCTTC	CAGTTCTCTG	ACAAAAAGTG	TCCAGAGGGC	3900
	TGTTTGCAAA	CACCTAGTGCA	CTTTGTAGCT	TTTCACCTCT	TGTCCCAGGG	AATCTAGGAG	3960
	AGATGAGGCC	CGTCAGAGTC	AAGAGATGTC	ATCCCCCAG	GGTCTCCAAG	GCATTTCCAC	4020
60	ACTATTGGTG	GCACCTGGAG	GACATGCACC	AAGCCTTGCC	AGAGCCACCA	GGAAGTGAGC	4080
	CCAGAGCATG	GCACATGAGC	ATCACCCGCT	GATGGTGGCC	TGCTGTGCCT	GGTGCCAAAC	4140
	GGGGCATCCC	GGCCCGTACC	CCTCCAGACA	GGAAGCATGG	GTTTGCCAC	AGACCTGTCTG	4200
	GGTGCTCCTG	TGAGTGGCCT	CCAGATGTCT	TTGTGCATAG	GCACAAGTGG	GCCAGGGCTG	4260
	GAGGGAGGTG	GGAAACCTCA	TCATCCGGTG	GGCCCTGCCA	ATCTTAACCC	AGAACCCTTA	4320
65	GGTATTCCTG	GCAGTAGCCA	TGACATTGGA	GCACCTTCTT	CTCCAGCCAG	AGGCTGACCT	4380
	GAGGCCCACT	GTCCTCAGAT	GACACCAACC	AGGACACCC	TAGGTGAGGG	GTGAGGGCCC	4440
	CCTTATGTGA	ACCTCTTGCC	TCTTCTTTTC	TCCCATCAGA	GTGGTTGGAT	GGAGCCATTG	4500
	GCCTCCTTTT	CTTACGCGGG	CCCTTCAACC	TCTCTGCACC	ATGTTGTCTG	GCTGAGGAGC	4560
	TACTAGAAAA	GCTGAGTGGA	GTCTCTTTTC	CAACAGGATG	ATGCATTGTC	TCAATTCTCA	4620
70	GGGCTGGAAT	GAGCCGGCTG	GTCCCCCAGA	AAGCTGGAGT	GGGGTACAGA	GTTTCACTTTT	4680
	CCTCTCTGTT	TACAGCTCCT	TGACAGTCCC	ACGCCCCTCT	GGAGTGGGAG	CTGGGAGTTA	4740
	GTGTTGGAGA	AGAAACAACA	AAAGCCAATT	AGAACCCTTA	TTTTTAAAAA	GTGCTTACTG	4800
	TGCACAGATA	CTCTTCAAGC	ACTGGACGTG	GATTCTCTCT	CTAGCCCTCA	GCACCCCTGC	4860
	GGTAGGAGTG	CCGCCTCTAC	CCACTTGTGA	TGGGGTACAG	AGGCACTTGC	TCTTCTGCAT	4920
75	GGTGTTCAT	AGGCTGGGAG	TTTTATTTAT	CTCTTCAAAC	TTTGTACAAG	AGCTCATGGC	4980
	TTGTCTTGGG	CTTTCGTCTAT	TAAACCAAAG	GAAATGGAAG	CCATTCCCTT	GTTGCTCTCC	5040
	TTAGTCTTGG	TCATCAGAAC	CTCACTTGGT	ACCATATAGA	TCAAAAGCTT	TGTAACCACA	5100
	GGAAAAAATA	AACTCTTCCA	TCCCTTAAAG	AAATAGATAG	TTTGTCCCTC	TCATGGGAAT	5160
	TGGGCTGTAT	GTATATGTTT	CTTCTCTCTT	AGAATTTAGA	GATACAGAG	TTCTACTTAG	5220
80	AACTTTTCAT	GGACACAATT	TCCACAACCT	TTTCAATGCT	GATGTAGAGC	TATTGGGAAA	5280
	GAACCTTCAA	ACTCAGGAAG	TTTGCAGAGA	GCAGACAGCT	AGAGATAACT	CGGGACCCAG	5340
	AGTTGGTCTGA	CAGATGTTAG	ATGTATCCTA	GCTTTTAGCC	ATAAACCACT	CAAAGATTCA	5400
	GCCCCAGAT	CCCACAGTCA	GAACCTGAATC	TGCGTTGTTG	GGAAGCCAGC	AGTGGCCTTG	5460
	GGAGGAAGC	CATGGCTGTG	GTTTCAAGAG	AGTGGGCTGG	CAAGCCACTT	CCGGGGAAAA	5520
85	CTCCTTCCGC	CCCAGGTTTC	TTCTTCTCTT	AAGGAGAGAT	TGTTCTCACC	AACCCGCTGC	5580
	CTTCATGCTG	CCTTCAAAGC	TAGATCATGT	TTGCCTTGCT	TAGAGAATTA	CTGCAAATCA	5640
	GCCCCAGTGC	TTGGCGATGC	ATTTACAGAT	TTCTAGGCCC	TCAGGGTTTT	GTAGAGTGTG	5700
	AGCCCTGGTG	GGCAGGGTTG	GGGGTCTGT	CTTCTGCTGG	ATGCTGCTTG	TAATCCATTT	5760



GGTGTACAGA ATCAACAATA AATAATATAC ATGTAT

Seq ID NO: 611 Protein sequence  
Protein Accession #: BAB84587.1

5  
10  
15

1	11	21	31	41	51	
MPLKHYLLLL	VGCQAWGAGL	AYHGCPSSECT	CSRASQVECT	GARIVAVPTP	LPWNAMSLQI	60
LNTHITELNE	SPFLNISALI	ALRIEKNELS	RITPGAFRNL	GSLRYLSLAN	NKLQVLPIGL	120
FQGLDSLES	LLSSNQLLQI	QPAHFSQCSN	LKELQLHGNH	LEYIPDGAFD	HLVGLTKLNL	180
GKNSLTHISP	RVFQHLGNLQ	VLRLYENRLT	DIPMGTFDGL	VNLQELALQQ	NQIGLLSPGL	240
FHNNHNLQRL	YLSNNHISQL	PPSIFMQLPQ	LNRLTLFGNS	LKELSLGIFG	PMPNLRELWL	300
YDNHISLSPD	NVFSNLRQLQ	VLILSRNQIS	FISPGAFNGL	TELRELSLHT	NALQDLDDGNV	360
FRMLANLQNI	SLQNNRLRQL	PGNIFANVNG	LMAIQLQNNQ	LENLPLGIFD	HLGKLCLELRL	420
YDNPWRCDSD	ILPLRNWLLL	NQPRLGTDTV	PVCFSPANVR	GQSLIIINVN	VAVPSVHVPE	480
VPSYPETPWY	PDTPSYPDTT	SVSSTTELTS	PVEDYTDLTT	IQVTDDRSVW	GMTQAQSGLA	540
IAAIVIGIVA	LACSLAACVG	CCCCCKRSQA	VLMQMKAPNE	C		

20

Seq ID NO: 612 DNA sequence  
Nucleic Acid Accession #: XM\_098151  
Coding sequence: 1..447

25  
30

1	11	21	31	41	51	
ATGATGCATT	TGCTCAATTC	TCAGGGCTGG	AATGAGCCGG	CTGGTCCCCC	AGAAAGCTGG	60
AGTGGGGTAC	AGAGTTCAGT	TTTCTCTCT	GTTTACAGCT	CCTTGACAGT	CCCACGCCCA	120
TCTGGAGTGG	GAGCTGGGAG	TCAGTGTGG	AGAAGAAACA	ACAAAAGCCA	ATTAGAACCA	180
CTATTTTTAA	AAAGTGCTTA	CTGTGCACAG	ATACTCTTCA	AGCACTGGAC	GTGGATTCTC	240
TCTCTAGCCC	TCAGCACCCC	TGCGGTAGGA	GTGCCGCCTC	TACCCACTTG	TGATGGGGTA	300
CAGAGGCACT	TGCTCTTCTG	CATGGTGTTC	AATAGGCTGG	GAGTTTTATT	TATCTCTTCA	360
AACTTTGTAC	AAGAGCTCAT	GGCTTGTCTT	GGGCTTTCGT	CATTAAACCA	AAGGAAATGG	420
AAGCCATTCC	CCTGTTGCTC	TCCTTAG				

35

Seq ID NO: 613 Protein sequence  
Protein Accession #: XP\_098151

40

1	11	21	31	41	51	
MMHLLNSQGW	NEPAGPPESW	SGVQSSVFLS	VYSSLTVPRP	SGVGAGSQCW	RRNNKSQLEP	60
LFLKSAYCAQ	ILFKHWTWIL	SLALSTPAVG	VPPLPTCDGV	QRHLLFCMVF	NRLGVLFISS	120
NFVQELMACL	GLSSLNQRKW	KPFPCCSP				

45

Seq ID NO: 614 DNA sequence  
Nucleic Acid Accession #: NM\_002658.1  
Coding sequence: 77..1372

50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
GTCCCCGAG	CGCCGTCGCG	CCCTCCTGCC	GCAGGCCACC	GAGGCCGCCG	CCGTCTAGCG	60
CCCCGACCTC	GCCACCATGA	GAGCCCTGCT	GGCGCGCCTG	CTTCTCTGCG	TCCTGGTCTG	120
GAGCGACTCC	AAAGGCAGCA	ATGAACCTCA	TCAAGTTCCA	TCGAAGTGTG	ACTGTCTAAA	180
TGGAGGAACA	TGTGTGTCCA	ACAAGTACTT	CTCCAACATT	CACTGGTGCA	ACTGCCCAAA	240
GAAATTCGGA	GGGCAGCACT	GTGAAATAGA	TAAGTCAAAA	ACCTGCTATG	AGGGGAATGG	300
TCACCTTTAC	CGAGGAAAGG	CCAGCACTGA	CACCATGGGC	CGGCCCTGCC	TGCCCTGGAA	360
CTCTGCCACT	GTCCTTCAGC	AAACGTACCA	TGCCCCACAGA	TCTGATGCTC	TTCAGCTGGG	420
CCTGGGGAAA	CATAATTACT	GCAGGAACCC	AGACAACCGG	AGGCGACCCCT	GGTGCTATGT	480
GCAGGTGGGC	CTAAAGCCGC	TTGTCCAAGA	GTGCATGGTG	CATGACTGCG	CAGATGGAAA	540
AAAGCCCTCC	TCTCCTCCAG	AAGAATTAAA	ATTTCACTGT	GGCCAAAAGA	CTCTGAGGCC	600
CCGCTTTAAG	ATTATTGGGG	GAGAATTCAC	CACCATCGAG	AACCAGCCCT	GGTTTGCGGC	660
CATCTACAGG	AGGCACCGGG	GGGGCTCTGT	CACCTACGTG	TGTGGAGGCA	GCCTCATCAG	720
CCCTTGCTGG	GTGATCAGCG	CCACACACTG	CTTCATTGAT	TACCCAAAGA	AGGAGGACTA	780
CATCGTCTAC	CTGGGTCTGCT	CAAGGCTTAA	CTCCAACACG	CAAGGGGAGA	TGAAGTTTGA	840
GGTGGAAAAC	CTCATCCTAC	ACAAGGACTA	CAGCGCTGAC	ACGCTTGCTC	ACCACAACGA	900
CATTGCCTTG	CTGAAGATCC	GTTCCAAGGA	GGGCAGGTGT	GCGCAGCCAT	CCCGGACTAT	960
ACAGACCATC	TGCCTGCCCT	CGATGCTATA	CGATCCCCAG	TTTGGCACAA	GCTGTGAGAT	1020
CACTGGCTTT	GGAAAAGAGA	ATTCTACCGA	CTATCTCTAT	CCGGAGCAGC	TGAAAATGAC	1080
TGTTGTGAAG	CTGATTTCCT	ACCGGGAGTG	TCAGCAGCCC	CACTACTACG	GCTCTGAAGT	1140
CACCACCAAA	ATGCTATGTG	CTGCTGACCC	CCAATGGAAA	ACAGATTCCCT	GCCAGGGAGA	1200
CTCAGGGGGA	CCCCTCGTCT	GTTCCCTCCA	AGGCCGCATG	ACTTTGACTG	GAATTGTGAG	1260
CTGGGGCCGT	GGATGTGCCC	TGAAGGACAA	GCCAGGCGTC	TACACGAGAG	TCTCACACTT	1320
CTTACCCTGG	ATCCGCAGTC	ACACCAATGG	AGCAATGGGC	CTGGCCCTCT	GAGGGTCCCC	1380
AGGGAGGAAA	CGGGCACCAC	CCGCTTTCTT	GCTGGTTGTC	ATTTTTCAG	TAGAGTCATC	1440
TCCATCAGCT	GTAAGAAGAG	ACTGGGAAGA	TAGGCTCTGC	ACAGATGGAT	TTGCCTGTGG	1500
CACCACCAGG	GTGAACGACA	ATAGCTTTAC	CCTCACGGAT	AGGCCTGGGT	GCTGGCTGCC	1560
CAGACCCTCT	GGCCAGGATG	GAGGGGTGGT	CCTGACTCAA	CATGTTACTG	ACCAGCAACT	1620
TGTCTTTTTT	TGGACTGAAG	CCTGCAGGAG	TTAAAAAGGG	CAGGGCATCT	CCTGTGCATG	1680
GGCTCGAAGG	GAGAGCCAGC	TCCCCGACC	GGTGGGCATT	TGTGAGGCC	ATGGTTGAGA	1740
AATGAATAAT	TTCCCAATTA	GGAAAGTGTA	GCAGCTGAGG	TCTCTTGAGG	GAGCTTAGCC	1800
AATGTGGGAG	CAGCGGTTTG	GGGAGCAGAG	ACACTAACGA	CTTCAGGGCA	GGGCTCTGAT	1860
ATTCCATGAA	TGTATCAGGA	AATATATATG	TGTGTGTATG	TTTGACACT	TGTTGTGTGG	1920
GCTGTGAGTG	TAAGTGTGAG	TAAGAGCTGG	TGTCTGATTG	TTAAGTCTAA	ATATTTCTCT	1980
AAACTGTGTG	GACTGTGATG	CCACACAGAG	TGGTCTTTCT	GGAGAGGTTA	TAGGTCACTC	2040
CTGGGGCCTC	TGGGTCCCC	CACGTGACAG	TGCCCTGGAA	TGTACTTATT	CTGCAGCATG	2100
ACCTGTGACC	AGCACTGTCT	CAGTTTCACT	TTACATAGA	TGTCCCTTTC	TTGGCCAGTT	2160
ATCCCTTCCT	TTTAGCCTAG	TTCATCCAAT	CCTCACTGGG	TGGGGTGAGG	ACCACTCCTT	2220
ACACTGAATA	TTTATATTTT	ACTATTTTTA	TTTATATTTT	TGTAATTTTA	AATAAAAGTG	2280
ATCAATAAAA	TGTGATTTTT	CTGA				



Seq ID NO: 615 Protein sequence  
Protein Accession #: NP\_002649.1

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
MRALLARLLL CVLVVSDSKG SNELHQVPSN CDCLNGGTCV SNKYFSNIHW CNCPKKFGGQ    60
HCEIDKSKTC YEGNGHFYRG KASTDTMGRP CLPWNSATVL QQTYHAHRSD ALQLGLGKHN    120
YCRNPDNRRR PWCYVQVGLK PLVQECMVHD CADGKKPSSP PEELKFQCGQ KTLRPRFKII    180
10    GGEFTTIENQ PWFAAIYRRH RGGSVTYVCG GSLISPCWVI SATHCFIDYP KKEDYIVYLG    240
RSRLNSNTQG EMKFEVENLI LHKDYSADTL AHHNDIALLK IRSKEGRCAQ PSRTIQTICL    300
PSMYNDPQFG TSCEITGFGK ENSTDYLYPE QLKMTVVKLI SHRECQPHY YGSEVTTKML    360
CAADPQWKTD SCQGDSSGGL VCSLQGRMTL TGIVSWGRGC ALKDKPGVYT RVSHFLPWIR    420
SHTKEENGLA L

```

Seq ID NO: 616 DNA sequence  
Nucleic Acid Accession #: NM\_024422.1  
Coding sequence: 202..2907

```

20    1      11      21      31      41      51
      |      |      |      |      |      |
CGCCAAAGGA AAAGCCCCTT GGATGAGAGG CAGGCGCTTC AGAGAAGCTA AGAAAAGCAC    60
CTCTCCGCGC GCCCCACCTC CTCCGCCTCG CGCTCCTCCT GAGCAGCGGG CCCAGACTGC    120
GCTCCGCGCG CGGCCCTCGC CCCGCGGAGC CCTCCTACCC CGGCCCGACG CTCGGCCCCG    180
25    GACCTGCCCC GAGCCCTCTC CATGGAGGCA GCCCGCCCCT CCGGCTCCTG GAACGGAGCC    240
CTCTGCCGGC TGCTCCTGCT GACCCTCGCG ATCTTAATAT TTGCCAGTGA TGCCTGCAAA    300
AATGTGACAT TACATGTTCC CTCCAAACTA GATCCGAGA AACTTGTTGG TAGAGTTAAC    360
CTGAAAGAGT GCTTTACAGC TGCAAATCTA ATTCATTCAA GTGATCCTGA CTTCCAAATT    420
TTGGAGGATG GTTCAGTCTA TACAACAAAT ACTATTCTAT TGTCTCTGGA GAAGAGAAGT    480
30    TTTACCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAATATT TGTCTTTTGT    540
GAGCATCAAA CAAAGGTCTT AAAGAAAAGA CATACTAAAG AAAAAGTTCT AAGGCGCGCC    600
AAGAGAAGAT GGGCTCCAAT TCCTTGTTCT ATGCTAGAAA ACTCCTTGGG TCCTTTTCCA    660
CTTTTCCTTC AACAGGTTCA ATCTGACACG GCCCAAAACT ATACCATATA CTATTCCATA    720
AGAGGTCTTG AGATTGACCA AGAACCTCGG AATTTATTTT ATGTGGAGAG AGACACTGGA    780
35    AACTTGTATT GTACTCGTCC TGTAGATCGT GAGCAGTATG AATCTTTTGA GATAATTGCC    840
TTTGCAACAA CTCCAGATGG GTATACTCCA GAACCTCCAC TGCCCCTAAT AATCAAAATA    900
GAGGATGAAA ATGATAACTA CCAATTTTTT ACAGAAAGAA CTTATACTTT TACAATTTTT    960
GAAAATTGCA GAGTGGGCAC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT    1020
GACACGATGC ACACACGCCT GAAGTACTCC ATCATTGGGC AGGTGCCACC ATCACCACC    1080
40    CTATTTTCTA TGCATCCAAC TACAGCGTGC ATCACCACAA CATCATCTCA GCTAGACAGA    1140
GAGTTAATTG ACAAGTACCA GTTGAAAATA AAAGTACAAG ACATGGATGG TCAGTATTTT    1200
GGTCTACAGA CAACTTCAAC TTGTATCATT AACATTGATG ATGTAAATGA CCACTTGCCA    1260
ACATTTACTC GTACTTCTTA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGGAAATC    1320
TTACGAGTTA CTGTTGAGGA TAAGGACTTA GTGAATACTG CTAAGTGGAG AGCTAATTAT    1380
45    ACCATTTTAA AGGGCAATGA AAATGGCAAT TTTAAATTTG TAACAGATGC CAAAACCAAT    1440
GAAGGAGTTC TTTGTGTAGT TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTTG    1500
CAAAATTGGT TAGTTAATGA AGCTCCATTT TCCAGAGAGG CTAGTCCAAG ATCAGCCATG    1560
AGCACAGCAA CAGTTACTGT TAATGTAGAA GATCAGGATG AGGGCCCTGA GTGTAACCCT    1620
50    CCAATACAGA CTGTTGCGAT GAAAGAAAAT GCAGAAAGTG GAACAACAAG CAATGGATAT    1680
AAAGCATATG ACCCAGAAAC AAGAAGTAGC AGTGGCATAA GGTATAAGAA ATTAAGTAT    1740
CCAACAGGGT GGGTCACCAT TGATGAAAAA ACAGGATCAA TCAAAGTTTT CAGAAGCCTG    1800
GATAGAGAGG CAGAGACCAT CAAAAATGGC ATATAATAA TTACAGTCCT TGCATCAGAC    1860
CAAGGAGGGA GAACATGTAC GGGGACACTG GGCAATTATC TTCAAGACGT GAATGATAAC    1920
AGCCCATTC AACTTAAAAA GACAGTGATC ATCTGCAAA CCAACATGTC ATCTGCGGAG    1980
55    ATTGTTGCGG TTGATCCTGA TGAGCCTATC CATGGCCAC CCTTTGACTT TAGTCTGGAG    2040
AGTTCTACTT CAGAAGTACA GAGAATGTGG AGACTGAAAG CAATTAATGA TACAGCAGCA    2100
CGTCTTTCCT ATCAGAAATGA TCCTCCATTT GGCTCATATG TAGTACCTAT AACAGTGAGA    2160
GATAGACTTG GCATGTCTAG TGTCACTTCA TTGGATGTTA CACTGTGTGA CTGCATTACC    2220
60    GAAAATGACT GCACATCGT TGTAGATCCA AGAATTGGCG GTGGAGGAGT ACAACTTGGA    2280
AAGTGGGCCA TCCTTGCAAT ATTGTTGGGC ATAGCATTGC TCTTTTGCAT CCTGTTTACG    2340
CTGGTCTGTG GGGCTTCTGG GACGTCTAAA CAACCAAAAG TAATTCCTGA TGATTTAGCC    2400
CAGCAGAACC TAATTGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGCG    2460
AATGGCTTCA CAACCCAAAC TGTGGGCGCT TCTGCTCAGG GAGTTTGTGG CACCGTGGGA    2520
65    TCAGGAATCA AAAACGGAGG TCAGGAGACC ATCGAAATGG TGAAAGGAGG ACACCAGACC    2580
TCGGAATCCT GCCGGGGGCG TGGCCACCAT CACACCTTGG ACTCCTGCAG GGGAGGACAC    2640
ACGGAGGTGG ACAACTGCAG ATACACTTAC TCGGAGTGGC ACAGTTTTAC TCAGCCCCGT    2700
CTTGGTGAAA AAGTGTATCT GTGTAATCAA GATGAAAATC ACAAGCATGC CCAAGACTAT    2760
70    GTCCTGACAT ATAAGTATGA AGGAAGAGGA TCGGTGGCTG GGTCTGTAGG TTGTTGCAGT    2820
GAACGACAAG AAGAAGATGG GCTTGAATTT TTGGATAATT TGGAGCCCAA ATTTAGGACA    2880
CTAGCAGAAG CATGCATGAA GAGATGAGTG GTTCTAATA AGTCTCTGAA AGCCAGTGGC    2940
TTTATGACTT TTAACAAAAA TTACAAACCA AGAATTTTTT AAAGCAGAAG ATGCTATTTG    3000
TGGGGGTTTT TCTCTCATTA TTTGGGATGA ATCTCTTTGG TCAAATGCAC ATTTACAGAG    3060
AGACACTATA AACAAGTACA CAAATTTTTT AATTTTTTACA TATTTTAAAT TACTTTATCT    3120
75    TCTATCCAAG GAGGTCTACA GAGAAATTAA AGTCTGCCTT ATTTGTTACA TTTGGGTATA    3180
ATGACAACAG CCAATTTATA GTGCAATAAA ATGTAATTAA TTCAAGTCCT TATTATAGAC    3240
TATTTGAAGC ACAACCTAAT GGAAATTTGT AGAGACCTTG CTTTAACATT ATCTCCAGTT    3300
AATTAAGTGT TCATGTGGTG CTTGGAAACT GTTGTTTTCC TGAACATCTA AAGTGTGTAG    3360
ACTGCATTCT TGCTATTATT TTATTCTTGT AATGTGACCT TTCACTGTG CAAAGGGAGA    3420
80    TTTCTAGCCA GGCATTGACT ATTACAATTT CATT

```

Seq ID NO: 617 Protein sequence  
Protein Accession #: NP\_077740.1

```

85    1      11      21      31      41      51
      |      |      |      |      |      |
MEAARPSGSW NGALCRLLLL TLAILIFASD ACKNVTLHVP SKLDAEKLVG RVNLKECFTA    60
ANLIHSSDPD FQILEDGSVY TTNTILLSSE KRSFTILLSN TENQEKKKIF VFLEHQTKVL    120

```



	KKRHTKEKVL	RRAKRRWAPI	PCSMLENSLG	PFPLFLQQVQ	SDTAQNYTIY	YSIRGPGVDQ	180
	EPRNLFYVER	DTGNLYCTR	VDREQYESFE	IIAFATTPDG	YTPLEPLPLI	IKIEDENDNY	240
	PIFTEETYTF	TIFENCRVGT	TVGQVCATDK	DEPDTMHTRL	KYSIIGQVPP	SPTLFMSHPT	300
5	TGVITTTSSQ	LDRELIDKYQ	LKIKVQDMDG	QYFGLQTTST	CIINIDDVND	HLPTFTRTSY	360
	VTSVEENTVD	VEILRVTVED	KDLVNTANWR	ANYTILKGNE	NGNFKIVTDA	KTNEGVLCCV	420
	KPLNYEEKQQ	MILQIGVVNE	APFSREASPR	SAMSTATVTV	NVEDQDEGPE	CNPPIQTVRM	480
	KENAEVGTTS	NGYKAYDPET	RSSSGIRYKK	LTDPTGWVTI	DENTGSIKVF	RSLDREAETI	540
	KNGIYNITVL	ASDQGGRTCT	GTLGIILQDV	NDNSPFIPKK	TVLICKPTMS	SAEIVAVDPD	600
10	EPIHGPPFDF	SLESTSEVQ	RMWRLKAIND	TAARLSYQND	PPFGSYVVP	TVRDRLGMS	660
	VTSLDVTLCD	CITENDCTHR	VDPRIGGGGV	QLGKWAILAI	LLGIALLF	LFTLVCGASG	720
	TSKQPKVIPD	DLAQQLNLI	NTEAPGDDKV	YSANGFTTQT	VGASAQGVCG	TVGSGIKNGG	780
	QETIEMVKGG	HQTSESCRGA	GHHHTLDSCR	GGHTEVDNCR	YTYSEWHSFT	QPRLGKVVYL	840
	CNQDENHKHA	QDYVLTYNYE	GRGSVAGSVG	CCSERQEEDG	LEFLDNLEPK	FRTLAEACMK	900
15	R						

Seq ID NO: 618 DNA sequence  
Nucleic Acid Accession #: NM\_004949.1  
Coding sequence: 202..2745

20	1	11	21	31	41	51	
	CGCCAAAGGA	AAAGCCCCTT	GGATGAGAGG	CAGGCGCTTC	AGAGAAGCTA	AGAAAAGCAC	60
	CTCTCCGCGC	GCCCCACCTC	CTCCGCCTCG	CGCTCCTCCT	GAGCAGCGGG	CCCAGACTGC	120
25	GCTCCGGCCG	CGGCCCTCGC	CCCGCGGAGC	CCTCCTACCC	CGGCCCGACG	CTCGGCCCGC	180
	GACCTGCCCC	GAGCCCTCTC	CATGGAGGCA	GCCCGCCCCC	CCGGCTCCTG	GAACGGAGCC	240
	CTCTGCCGGC	TGCTCCTGCT	GACCCTCGCG	ATCTTAATAT	TGCGCAGTGA	TGCCTGCAAA	300
	AATGTGACAT	TACATGTTCC	CTCCAAACTA	GATGCCGAGA	AACTTGTTGG	TAGAGTTAAC	360
	CTGAAAGAGT	GCTTTACAGC	TGCAAACTCT	ATTCAATCAA	GTGATCCTGA	CTTCCAAATT	420
30	TTGGAGGATG	GTTCACTCTA	TACAACAAAT	ACTATTCTAT	TGTCCTCGGA	GAAGAGAAGT	480
	TTTACCATAT	TACTTTCCAA	CACTGAGAAC	CAAGAAAAGA	AGAAAATATT	TGTCTTTTTG	540
	GAGCATCAAA	CAAAGGTCC	AAAGAAAAGA	CATACTAAAG	AAAAAGTTCT	AAGGCGCGCC	600
	AAGAGAAGAT	GGGCTCCAA	TCCTTGTTCC	ATGCTAGAAA	ACTCCTTGGG	TCCTTTTCCA	660
	CTTTTCCTTC	AACAGGTTCA	ATCTGACACG	GCCCAAAACT	ATACCATATA	CTATTCCATA	720
35	AGAGGTCCCT	GAGTTGACCA	AGAACCTCGG	AATTTATTTT	ATGTGGAGAG	AGACACTGGA	780
	AACTTGATAT	GTACTCGTCC	TGTAGATCGT	GAGCAGTATG	AATCTTTTGA	GATAATTGCC	840
	TTTGCAACAA	CTCCAGATGG	GTATACTCCA	GAATTCCAC	TGCCCCTAAT	AATCAAAATA	900
	GAGGATGAAA	ATGATAACTA	CCCAATTTTT	ACAGAAGAAA	CTTATACTTT	TACAATTTTT	960
	GAAAATTGCA	GAGTGGGCAC	TACTGTGGGG	CAAGTGTGTG	CTACTGACAA	AGATGAGCCT	1020
40	GACACGATGC	ACACACGCC	GAAGTACTCC	ATCATTTGGG	AGGTGCCACC	ATCACCACC	1080
	CTATTTTCTA	TGCATCCAAC	TACAGGCGTG	ATCACCACAA	CATCATCTCA	GCTAGACAGA	1140
	GAGTTAATTG	ACAAGTACCA	GTTGAAAATA	AAAGTACAAG	ACATGGATGG	TCAGTATTTT	1200
	GGTCTACAGA	CAACTTCAAC	TTGTATCATT	AACATTGATG	ATGTAAATGA	CCACTTGCCA	1260
	ACATTTACTC	GTACTTCTTA	TGTGACATCA	GTGGAAGAAA	ATACAGTTGA	TGTGGAAATC	1320
45	TTACGAGTTA	CTGTTGAGGA	TAAGGACTTA	GTGAATAC	CTAACTGGAG	AGCTAATTAT	1380
	ACCATTTTAA	AGGGCAATGA	AAATGGCAAT	TTTAAATTTG	TAACAGATGC	CAAAACCAAT	1440
	GAAGGATGTT	TTTGTGTAGT	TAAGCTTTTG	AATTATGAAG	AAAAGCAACA	GATGATCTTG	1500
	CAAATTGGTG	TAGTTAATGA	AGCTCCATTT	TCCAGAGAGG	CTAGTCCAAG	ATCAGCCATG	1560
	AGCACAGCAA	CAGTTACTGT	TAATGTAGAA	GATCAGGATG	AGGGCCCTGA	GTGTAACCCT	1620
50	CCAATACAGA	CTGTTTCGAT	GAAAGAAAAT	GCAGAAGTGG	GAACAACAAG	CAATGGATAT	1680
	AAAGCATATG	ACCCAGAAAC	AAGAAAGTAG	AGTGGCATAA	GGTATAAGAA	ATTAAGTATG	1740
	CCAACAGGGT	GGGTCAACAT	TGATGAAAA	ACAGGATCAA	TCAAAGTTTT	CAGAAGCCTG	1800
	GATAGAGAGG	CAGAGACCAT	CAAAAATGGC	ATATATAATA	TTACAGTCCT	TGCATCAGAC	1860
	CAAGGAGGGA	GAACATGTAC	GGGGACACTG	GGCATTATAC	TTCAAGACGT	GAATGATAAC	1920
55	AGCCCATTCA	TACCTAAAA	GACAGTGATC	ATCTGCAAA	CCACCATGTC	ATCTGCGGAG	1980
	ATTGTTGCGG	TTGATCCTGA	TGAGCCTATC	CATGGCCAC	CCTTTGACTT	TAGTCTGGAG	2040
	AGTTCTACTT	CAGAAGTACA	GAGAATGTGG	AGACTGAAAG	CAATTAATGA	TACAGCAGCA	2100
	CGTCTTTCCT	ATCAGAATGA	TCCTCCATTT	GGCTCATATG	TAGTACCTAT	AACAGTGAGA	2160
	GATAGACTTG	GCATGTCTAG	TGTCACTTCA	TTGGATGTTA	CACTGTGTGA	CTGCATTACC	2220
60	GAAAATGACT	GCACACATCG	TGTAGATCCA	AGGATGGCG	GTGGAGGAGT	ACAACCTTGA	2280
	AAGTGGGCCA	TCCTTGCAAT	ATTGTTGGGC	ATAGCATTGC	TCTTTTGCAT	CCTGTTTACG	2340
	CTGGTCTGTG	GGGCTTCTGG	GACGTCTAAA	CAACCAAAAG	TAATTCCTGA	TGATTTAGCC	2400
	CAGCAGAACC	TAATTGTATC	AAACACAGAA	GCTCCTGGAG	ATGACAAAGT	GTATTCTGCG	2460
	AATGGCTTCA	CAACCCAAAC	TGTGGGCGCT	TCTGCTCAGG	GAGTTTGTGG	CACCGTGGGA	2520
65	TCAGGAATCA	AAAACGGAGG	TCAGGACACC	ATCAGAAATG	TGAAAGGAGG	ACACCAGACC	2580
	TCGGAATCCT	GCCGGGGGCG	TGGCCACCAT	CACACCTTGG	ACTCCTGCAG	GGGAGGACAC	2640
	ACGGAGGTGG	ACAACCTGCAG	ATACACTTAC	TCGGAGTGGC	ACAGTTTAC	TCAGCCCCGT	2700
	CTTGGTGAAG	AATCCATTAG	AGGACACACT	CTGATTAAAA	ATTAAACAAT	GAAAGAAAGT	2760
	GTATCTGTGT	AATCAAGATG	AAAATCACAA	GCATGCCCAA	GACTATGTCC	TGACATATAA	2820
70	CTATGAAGGA	AGAGGATCGG	TGGCTGGGTC	TGTAGGTTGT	TGCAGTGAAC	GACAAGAAGA	2880
	AGATGGGCTT	GAATTTTGG	ATAATTTGGA	GCCCAAATTT	AGGACACTAG	CAGAAGCATG	2940
	CATGAAGAGA	TGAGTGTGTT	CTGAAAGGCC	TCTGAAAGCC	AGTGGCTTTA	TGACTTTTAA	3000
	AAAAAATTAC	AAACCAAGAA	TTTTTTAAAG	CAGAAGATGC	TATTTGTGGG	GGTTTTTCTC	3060
	TCATTATTTG	GATGGAATCT	CTTTGGTCAA	ATGCACATTT	ACAGAGAGAC	ACTATAAACA	3120
75	AGTACACAAA	TTTTTCAATT	TTTACATATT	TTTAAATTAC	TTATCTTCTA	TCCAAGGAGG	3180
	TCTACAGAGA	AATTAAAGTC	TGCCTTATTT	GTTACATTTG	GGTATAATGA	CAACAGCCAA	3240
	TTTATAGTGC	AATAAAATGT	AATTAATTCA	AGTCCTTATT	ATAGACTATT	TGAAGCACAA	3300
	CCTAATGGAA	AATTGTAGAG	ACCTTGCTTT	AACATTATCT	CCAGTTAATT	AAGTGTTTAT	3360
	GTGGTGCTTG	GAAACTGTTG	TTTTCCTGAA	CATCTAAAGT	GTGTAGACTG	CATTCTTGCT	3420
80	ATTATTTTAT	TCTTGTAATG	TGACCTTTTC	ACTGTGCAAA	GGGAGATTTC	TAGCCAGGCA	3480
	TTGACTATTA	CAATTTTATT					

Seq ID NO: 619 Protein sequence  
Protein Accession #: NP\_004940.1

85	1	11	21	31	41	51	
	MEARPSGSW	NGALCRLLLL	TLAILIFASD	ACKNVTLHVP	SKLDAEKLVG	RVNLKECF	60



ANLIHSSDPD FQILEDGSVY TTNTILLSSE KRSFTILLSN TENQEKKKIF VFLEHQTKVL 120  
KKRHTKEKVL RRAKRRWAPI PCSMLENSLG PFPLFLQQVQ SDTAQNYTIY YSIRGPGVDQ 180  
EPRNLFYVER DTGNLYCTRP VDREQYESFE IIAFATTPDG YTPELPLPLI IKIEDENDNY 240  
5 PIFTEETYTF TIFENCRVGT TVGQVCATDK DEPDTHMTRL KYSIIGQVPP SPTLFMSHPT 300  
TGVITTTSSQ LDRELIDKYQ LKIKVQDMDG QYFGLQTTST CIINIDDVND HLPFTFRTSY 360  
VTSVEENTVD VEILRVTVED KDLVNTANWR ANYTILKGNE NGNFKIVTDA KTNEGVL CVV 420  
KPLNYEEKQQ MILQIGVVNE APFSREASPR SAMSTATVTV NVEDQDEGPE CNPPIQTVM 480  
KENAEVGTTS NGYKAYDPET RSSSGIRYKK LTDPTGWVTI DENTGSIKVF RSLDREAETI 540  
10 KNGIYNITVL ASDQGGRTCT GTLGIILQDV NDNSPFIPKK TVIICKPTMS SAEIVAVDPD 600  
EPIHGPPPDF SLESSTSEVQ RMWRLKAIND TAARLSYQND PPFGSYVVP I TVRDLGMSS 660  
VTSLDVTLCD CITENDCTHR VDPRIGGGGV QLKGWAILAI LLGIALLCFI LFTLVCGASG 720  
TSKQPKVIPD DLAQONLIVS NTEAPGDDKV YSANGFTTQT VGASAQGVCG TVGSGIKNGG 780  
QETIEMVKGK HQTSESCRGA GHHTLDSR GGHTEVDNCR YTYSEWSFT QPRLGEESIR 840  
GHTLIKN

Seq ID NO: 620 DNA sequence  
Nucleic Acid Accession #: NM\_032545.1  
Coding sequence: 46..718

20 1 11 21 31 41 51  
| | | | | |  
AAACTGATCT TCAATGCACT AAGAGAAGGA GACTCTCAAA CCAAAAATGA CCTGGAGGCA 60  
CCATGTCAGG CTTCTGTTTA CGGTCAGTTT GGCATTACAG ATCATCAATT TGGGAAACAG 120  
25 CTATCAAAGA GAGAAACATA ACGGCGGTAG AGAGGAAGTC ACCAAGGTTG CCACTCAGAA 180  
GCACCGACAG TCACCGCTCA ACTGGACCTC CAGTCATTTC GGAGAGGTGA CTGGGAGCGC 240  
CGAGGGCTGG GGGCCGGAGG AGCCGCTCCC TACTCCCAG GCTTTCGAG AGGGTGCCTC 300  
CGCGCGGCCG CGCTGCTGCA GGAACGGCGG TACCTGCGTG CTGGGCAGCT TCTGCGTGTG 360  
CCCGGCCAC TTCACCGGCC GCTACTGCGA GCATGACCAG AGGCGCAGTG AATGCGGCGC 420  
30 CCTGGAGCAC GGAGCCTGGA CCCTCCGCGC CTGCCACCTC TGCAGGTGCA TCTTCGGGGC 480  
CCTGCACTGC CTCCCCCTCC AGACGCCTGA CCGCTGTGAC CCGAAAGACT TCCTGGCCTC 540  
CCACGCTCAC GGGCCGAGCG CCGGGGGCGC GCCCAGCCTG CTACTCTTGC TGCCCTGCGC 600  
ACTCCTGCAC CGCCTCCTGC GCCCGGATGC GCCCGCGCAC CCTCGGTCCC TGGTCCCTTC 660  
CGTCTCCAG CGGGAGCGGC GCCCTGCGG AAGCCGGGCA CTTGGGCATC GCCTTTAATT 720  
35 TTCTATGTTG TAAATAATAG ATGTGTTTGG TTTACCGTAA GCTGAAGCAC TGGGTGAATA 780  
TTTTTATTGG GTAATAAATA TTTTCATGAA AGCGCCAAAA AAAAAAAAAA AAAAAAAAAA 840  
AAAAAA

Seq ID NO: 621 Protein sequence  
Protein Accession #: NP\_115934.1

40 1 11 21 31 41 51  
| | | | | |  
MTWRHHVRL FTVSLALQII NLGNSYQREK HNGGREEVTK VATQKHRQSP LNWTSSHFG 60  
45 VTGSAEGWGP EEPLPYSAF GEGASARPRC CRNGGTCVLG SFCVCPAHFT GRYCEHDQRR 120  
SECGALEHGA WTLRACHLCR CIFGALHCLP LQTPDRCDPK DFLASHANGP SAGGAPSLLL 180  
LLPCALLHRL LRPDAPAHPR SLVPSVLQRE RRPCGRPLG HRL

Seq ID NO: 622 DNA sequence  
Nucleic Acid Accession #: FGENESH predicted  
Coding sequence: 1..390

50 1 11 21 31 41 51  
| | | | | |  
55 ATGAGGTTCA GTGTCTCAGG CATGAGGACC GACTACCCCA GGAGTGTGCT GGCTCCTGCT 60  
TATGTGTGAG TCTGTCTCCT CCTCTTGTGT CCAAGGGAAG TCATCGCTCC CGCTGGCTCA 120  
GAACCATGGC TGTGCCAGCC GGCACCCAGG TGTGGAGACA AGATCTACAA CCCCTTGGAG 180  
CAGTGTCTGT ACAATGACGC CATCGTGTCC CTGAGCGAGA CCCGCCAATG TGGTCCCCC 240  
TGCACCTTCT GGCCCTGCTT TGAGCTCTGC TGTCTTGATT CCTTTGGCCT CACAAACGAT 300  
60 TTTGTTGTGA AGCTGAAGGT TCAGGGTGTG AATTCCCAGT GCCACTCATC TCCCATCTCC 360  
AGTAAATGTG AAAGAGGCCG GATATGTTAG

Seq ID NO: 623 Protein sequence  
Protein Accession #: FGENESH predicted

65 1 11 21 31 41 51  
| | | | | |  
MRFVSGMRT DYPRSVLAPA YVSVCLLLC PREVIAPAGS EPWLCQPAPR CGDKIYNPLE 60  
70 QCCYNDAIVS LSETRQCGPP CTFWPCFELC CLDSFGLTND FVVKLVQGV NSQCHSSPIS 120  
SKCERGRIC

Seq ID NO: 624 DNA sequence  
Nucleic Acid Accession #: M18728.1  
Coding sequence: 51..1085

75 1 11 21 31 41 51  
| | | | | |  
GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60  
CCTCAGCCCC TCCCTGCAGA TTGCATGTCC CCTGGAAGGA GGTCCTGCTC ACAGCCTCAC 120  
80 TTCTAACCTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180  
ATGTCGCAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAA CCTGCCCCAG AATCGTATTG 240  
GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300  
TAGGAACTCA ACAAGCTACC CCAGGGCCCC CATAAGTGG TCGAGAGACA ATATACCCCA 360  
ATGCATCCCT GCTGATCCAG AACGTACCC AGAATGACAC AGGATTCTAT ACCCTACAAG 420  
TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGGAGC 480  
85 TGCCCAAGCC CTCCATCTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540  
CCTTCACCTG TGAACCTGAG GTTCAGAACA CAACCTACCT GTGGTGGGTA AATGGTCAGA 600  
GCCTCCCGGT CAGTCCCAGG CTGCAGCTGT CCAATGGCAA CATGACCCTC ACTCTACTCA 660



35 Seq ID NO: 625 Protein sequence  
Protein Accession #: AAA59907.1

Seq ID NO: 626 DNA sequence  
Nucleic Acid Accession #: M18728.1  
Coding sequence: 1355..1657

427



TCTCACCTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220  
GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAAC CAATTTAAAA AAAAAAAGA 2280  
ACACAGGAGA TTCCAGTCTA CTTGAGTTAG CATAATACAG AAGTCCCCTC TACTTTAACT 2340  
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCAA TGTATTTATT TCTGTGGTTC 2400  
TGTTTCCTTG TTCCAATTTG ACAAACCCA CTGTTCTTGT ATTGTATTGC CCAGGGGGAG 2460  
CTATCACTGT ACTTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAAT AAAAGCCAAT 2520  
TAGCTCTATA ACT

Seq ID NO: 627 Protein sequence  
Protein Accession #: AAA59908.1

1 11 21 31 41 51  
| | | | | |  
MDSFSQDVKT RLLIMIRLLP PFNLSLLMPA SFAWQDDAVI SISQEVASEG NLTECQIYLV 60  
NPNVLHKIRD PLVHPVTDIS SIFNTAVCSN VQWSFSELDL

Seq ID NO: 628 DNA sequence  
Nucleic Acid Accession #: M18728.1  
Coding sequence: 2370..2501

1 11 21 31 41 51  
| | | | | |  
GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60  
CCTCAGCCCC TCCCTGCAGA TTGCATGTCC CCTGGAAGGA GGTCTGCTC ACAGCCTCAC 120  
TTCTAACCTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180  
ATGTCGCAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAA CCTGCCCCAG AATCGTATTG 240  
GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300  
TAGGAACCTA ACAAGCTACC CCAGGGCCCG CATACTGTGG TCGAGAGACA ATATACCCCA 360  
ATGCATCCCT GCTGATCCAG AACGTCACCC AGAATGACAC AGGATTCTAT ACCCTACAAG 420  
TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGGAGC 480  
TGCCCCAAGCC CTCCATCTCC AGCAACAACCT CCAACCCCGT GGAGGACAAG GATGCTGTGG 540  
CCTTCACCTG TGAACCTGAG GTTCAGAACA CAACCTACCT GTGGTGGGTA AATGGTCAGA 600  
GCCTCCCGGT CAGTCCCAGG CTGACCTGT CCAATGGCAA CATGACCTC ACTCTACTCA 660  
GCGTCAAAAG GAACGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720  
ACCGCAGTGA CCCAGTCACC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCC 780  
CCTCAAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCTGCCCAC GCAGCCTCTA 840  
ACCCACCTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900  
TCTTTATCCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960  
CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGCTCCTG 1020  
TCCTCTCAGC TGTGGCCACC GTCGGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080  
TATAGCAGCC CTGGTGTATT TTCGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCCT 1140  
GAATTCCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGTCTG 1200  
CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAATC AATGAAAATT TAAAGGGAAA 1260  
ACCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACTAGAGA CAGTCAAAT 1320  
GCAAACCATG GTGAGAAATT GACGACTTCA CACTATTGGC AGCTTTTCCC AAGATGTCAA 1380  
AACAAGACTC CTATCATGA TAAGGCTCTT ACCCCCTTTT AATTTGTCCT TGCTTATGCC 1440  
TGCTCTTTT GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500  
GGGTAACCTA ACAGAGTGTC AGATCTATCT TGTCAATCCC AACGTTTAC ATAAAATAAG 1560  
AGATCCTTTA GTGCACCCAG TGAATGACAT TAGCAGCATC TTTAACACAG CCGTGTGTTC 1620  
AAATGTACAG TGGTCTTTT CAGAGTTGGA CTCTAGACT CACCTGTTCT CACTCCCTGT 1680  
TTTAATTCAA CCCAGCCATG CAATGCCAAA TAATAGAATT GCTCCCTACC AGCTGAACAG 1740  
GGAGGAGTCT GTGAGTTTCT TGACACTTGT TGTGAAACAT GGCTAAATAC AATGGGTATC 1800  
GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860  
CTGACTCATT CTTTATTCTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920  
CTCTTGGTAT TACCTCCTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980  
CTCTAAAAGC TTTAAATGTC TGCATGCAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040  
GGCTGGAATT ACAAACCTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100  
ATAAAAGCCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTGGTCACAC 2160  
TCTCACCTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220  
GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAAC CAATTTAAAA AAAAAAAGA 2280  
ACACAGGAGA TTCCAGTCTA CTTGAGTTAG CATAATACAG AAGTCCCCTC TACTTTAACT 2340  
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCAA TGTATTTATT TCTGTGGTTC 2400  
TGTTTCCTTG TTCCAATTTG ACAAACCCA CTGTTCTTGT ATTGTATTGC CCAGGGGGAG 2460  
CTATCACTGT ACTTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAAT AAAAGCCAAT 2520  
TAGCTCTATA ACT

Seq ID NO: 629 Protein sequence  
Protein Accession #: AAA59909.1

1 11 21 31 41 51  
| | | | | |  
MLTNVFISVV LFPCSNLTKP TVLVLYCPGG AITVLVEWCC FNS

Seq ID NO: 630 DNA sequence  
Nucleic Acid Accession #: NM\_016639.1  
Coding sequence: 40..429

1 11 21 31 41 51  
| | | | | |  
GCGGCGGGCG CAGACAGCGG CGGGCGCAGG ACGTGCACTA TGGCTCGGGG CTCGCTGCGC 60  
CGGTTGCTGC GGTCTCTCGT GCTGGGGCTC TGGCTGGCGT TGCTGCGCTC CGTGGCCGGG 120  
GAGCAAGCGC CAGGCACCGC CCCCTGCTCC CGCGGACGCT CCTGGAGCGC GGACCTGGAC 180  
AAGTGCATGG ACTGCGCGTC TTGCGGGCTG CGACCGCACA GCGACTTCTG CCTGGGCTGC 240  
GCTGCAGCAC CTCCTGCCCC CTTCCGGCTG CTTTGGCCCA TCCTTGGGGG CGCTCTGAGC 300  
CTGACCTTCG TGCTGGGGCT GCTTTCTGGC TTTTGGTCTT GGAGACGATG CCGCAGGAGA 360  
GAGAAGTTCA CCACCCCAT AGAGGAGACC GGCGGAGAGG GCTGCCCAGC TGTGGCGCTG 420



ATCCAGTGAC AATGTGCCCC CTGCCAGCCG GGGCTCGCCC ACTCATCATT CATTCATCCA 480  
 TTCTAGAGCC AGTCTCTGCC TCCCAGACGC GGCGGGAGCC AAGCTCCTCC AACCACAAGG 540  
 GGGGTGGGGG GCGGTGAATC ACCTCTGAGG CCTGGGCCCA GGGTTCAGGG GAACCTTCCA 600  
 AGGTGTCTGG TTGCCCTGCC TCTGGCTCCA GAACAGAAAG GGAGCCTCAC GCTGGCTCAC 660  
 ACAAACAGC TGACACTGAC TAAGGAAGT CAGCATTTGC ACAGGGGAGG GGGGTGCCCT 720  
 CCTTCCTTAG GACCTGGGGG CCAGGCTGAC TTGGGGGGCA GACTTGACAC TAGGCCCCAC 780  
 TCACTCAGAT GTCCTGAAAT TCCACCACGG GGGTCACCCT GGGGGGTTAG GGACCTATTT 840  
 TTAACACTAG GGGCTGGCCC ACTAGGAGGG CTGGCCCTAA GATACAGACC CCCCCAATC 900  
 CCCAAAGCGG GGAGGAGATA TTTATTTTGG GGAGAGTTG GAGGGGAGGG AGAATTTATT 960  
 AATAAAGAA TCTTTAACTT TAAAAAATA AAAAAA

Seq ID NO: 631 Protein sequence  
 Protein Accession #: NP\_057723.1

1 11 21 31 41 51  
 | | | | |  
 MARGSLRRLL RLLVLGLWLA LLRSVAGEQA PGTAPCSRGS SWSADLDKCM DCASCRARPH 60  
 SDFCLGCAA PPAPFRLLWP ILGGALSLTF VLGLLSGFLV WRRCRREKEF TTPIEETGGE 120  
 GCPAVALIQ

Seq ID NO: 632 DNA sequence  
 Nucleic Acid Accession #: NM\_003816.1  
 Coding sequence: 79..2538

1 11 21 31 41 51  
 | | | | |  
 CGGCAGGGTT GGAAATGAT GGAAGAGGCG GAGGTGGAGG CGACCGAGTG CTGAGAGGAA 60  
 CCTGCGGAAT CGGCCGAGAT GGGGTCTGGC GCGCGCTTTC CCTCGGGGAC CCTTCGTGTC 120  
 CGGTGGTTGC TGTGCTTGG CCTGGTGGGC CCAGTCCTCG GTGCGGCGCG GCCAGGCTTT 180  
 CAACAGACCT CACATCTTTC TTCTTATGAA ATTATAACTC CTTGGAGATT AACTAGAGAA 240  
 AGAAGAGAAG CCCCTAGGCC CTATTCAAAA CAAGTATCTT ATGTTATTCA GGCTGAAGGA 300  
 AAAGAGCATA TTATTCACCT GGAAAGGAAC AAAGACCTTT TGCCTGAAGA TTTTGTGGTT 360  
 TATACTTACA ACAAGGAAGG GACTTTAATC ACTGACCATC CCAATATACA GAATCATTGT 420  
 CATTATCGGG GCTATGTGGA GGGAGTTTCAT AATTCATCCA TTGCTCTTAG CGACTGTTTT 480  
 GGACTCAGAG GATTGCTGCA TTTAGAGAAT GCGAGTTATG GGATTGAACC CCTGCAGAAC 540  
 AGCTCTCATT TTGAGCACAT CATTTATCGA ATGGATGATG TCTACAAAGA GCCTCTGAAA 600  
 TGTGGAGTTT CCAACAAGGA TATAGAGAAA GAAACTGCAA AGGATGAAGA GGAAGAGCCT 660  
 CCCAGCATGA CTCAGCTACT TCGAAGAAGA AGAGCTGTCT TGCCACAGAC CCGGTATGTG 720  
 GAGCTGTTCA TTGTCGTAGA CAAGGAAAGG TATGACATGA TGGGAAGAAA TCAGACTGCT 780  
 GTGAGAGAAG AGATGATTCT CCTGGCAAAC TACTTGGATA GTATGTATAT TATGTTAAAT 840  
 ATTCGAATTG TGCTAGTTGG ACTGGAGATT TGGACCAATG GAAACCTGAT CAACATAGTT 900  
 GGGGTGCTG GTGATGTGCT GGGGAACTTC GTGCAGTGGC GGGAAAAGTT TCTTATCACA 960  
 CGTCGGAGAC ATGACAGTGC ACAGCTAGTT CTAAAGAAAG GTTTTGGTGG AACTGCAGGA 1020  
 ATGGCATTTG TGGGAACAGT GTGTTCAAGG AGCCACGCAG GCGGGATTAA TGTGTTTGGG 1080  
 CAAATCACTG TGGAGACATT GTCTTCCATT GTTGCTCATG AATTGGGTCA TAATCTTGGA 1140  
 ATGAATCAGC ATGATGGGAG AGATTGTTCC TGTGGAGCAA AGAGCTGCAT CATGAATTCA 1200  
 GGAGCATCGG GTTCCAGAAA CTTTAGCAGT TGCAGTGCAG AGGACTTTGA GAAGTTAACT 1260  
 TTAAATAAAG GAGGAAACTG CCTTCTTAAT ATTCCAAAGC CTGATGAAGC CTATAGTGCT 1320  
 CCCTCCTGTG GTAATAAGTT GGTGGACGCT GGGGAAGAGT GTGACTGTGG TACTCCAAAG 1380  
 GAATGTGAAT TGGACCTTGG CTGCGAAGGA AGTACCTGTA AGCTTAAATC ATTTGCTGAG 1440  
 TGTGCATATG GTGACTGTTG TAAAGACTGT CGGTTCCTTC CAGGAGGTAC TTTATGCCGA 1500  
 GGAAAAATCA GTGAGTGTGA TGTTCAGAG TACTGCAATG GTTCTTCTCA GTTCTGTGAG 1560  
 CCAGATGTTT TTATTCAGAA TGGATATCCT TGCCAGAATA ACAAAGCCTA TTGCTACAAC 1620  
 GGCATGTGCC AGTATTATGA TGCTCAATGT CAAGTCATCT TTGGCTCAA AGCCAAGGCT 1680  
 GCCCCCAAAG ATTGTTTCAT TGAAGTGAAT TCTAAAGGTG ACAGATTTGG CAATTGTGGT 1740  
 TTCTCTGGCA ATGAATACAA GAAGTGTGCC ACTGGGAATG CTTTGTGTGG AAAGCTTCAG 1800  
 TGTGAGAATG TACAAGAGAT ACCTGTATTT GGAATTGTGC CTGCTATTAT TCAAACGCCCT 1860  
 AGTCGAGGCA CCAAATGTTG GGGTGTGGAT TTCCAGCTAG GATCAGATGT TCCAGATCCT 1920  
 GGGATGTTTA ACGAAGGCAC AAAATGTGGT GCTGGAAAGA TCTGTAGAAA CTTCCAGTGT 1980  
 GTAGATGCTT CTGTTCTGAA TTATGACTGT GATGTTTCTA AAAAGTGTCA TGGACATGGG 2040  
 GTATGTAATA GCAATAAGAA TTGTCACTGT GAAATGGCT GGGCTCCCC AAATTGTGAG 2100  
 ACTAAAGGAT ACGGAGGAAG TGTGGACAGT GGACCTACAT ACAATGAAAT GAATACTGCA 2160  
 TTGAGGGACG GACTTCTGGT CTTCTTCTTC CTAATTGTTC CCCTTATTGT CTGTGCTATT 2220  
 TTTATCTTCA TCAAGAGGGA TCAACTGTGG AGAAGCTACT TCAGAAAGAA GAGATCACAA 2280  
 ACATATGAGT CAGATGGCAA AAATCAAGCA AACCCTTCTA GACAGCCGGG GAGTGTTCCT 2340  
 CGACATGTTT CTCCAGTGAC ACCTCCAGAG GAAGTTCCTA TATATGCAA CAGATTTGCA 2400  
 GTACCAACCT ATGCAGCCAA GCAACCTCAG CAGTTCCTAT CAAGGCCACC TCCACCACAA 2460  
 CCGAAAGTAT CATCTCAGGG AAACCTAATT CCTGCCCGTC CTGCTCCTGC ACCTCCTTTA 2520  
 TATAGTTCCC TCACTTGATT TTTTAAACCT TCTTTTTCGA AATGTCTTCA GGGAACTGAG 2580  
 CTAATACTTT TTTTTTTTCT TGATGTTTTT TTGAAAAGCC TTTCTGTTGC AACTATGAAT 2640  
 GAAAAACAAA CACCACAAA CAGACTTCAC TAACACAGAA AAACAGAAAC TGAGTGTGAG 2700  
 AGTTGTGAAA TACAAGGAAA TGCAGTAAAG CCAGGGAATT TACAATAACA TTTCCGTTTC 2760  
 CATCATTGAA TAAGTCTTAT TCAGTCATCG GTGAGGTTAA TGCATAATC ATGGATTTTT 2820  
 TGAACATGTT ATTGCAGTGA TTCTCAAATT AACTGTATTG GTGTAAGATT TTTGTCAATTA 2880  
 AGTGTTTAAG TGTATTCTG AATTTTCTAC CTTAGTTATC ATTAATGTAG TTCTCATTG 2940  
 AACATGTGAT AATCTAATAC CTGTGAAAAA TGACTAATCA GCTGCCAATA ATATCTAATA 3000  
 TTTTTCATCA TGCACGAATT AATAATCATC ATACTTAGA ATCTTGTCTG TCACTCACTA 3060  
 CATGAATAAG CAAATATTGT CTTCAAAAGA ATGCACAAGA ACCACAATTA AGATGTCATA 3120  
 TTATTTTGAA AGTACAAAAT ATACTAAAAG AGTGTGTGTG TATTCACGCA GTTACTCGCT 3180  
 TCCATTTTTA TGACCTTTCA ACTATAGGTA ATAACCTTA GAGAAATTAA TTTAATATTA 3240  
 GAATTTCTAT TATGAATCAT GTGAAAGCAT GACATTCTGT CACAATAGCA CTATTTTAAA 3300  
 TAAATTATAA GCTTTAAGGT ACGAAGTATT TAATAGATCT AATCAAATAT GTTGATTTCAT 3360  
 GGCTATAATA AAGCAGGAGC AATTATAAAA TCTTCAATCA ATTGAACCTT TACAAAACCA 3420  
 CTTGAGAAAT TCATGAGCAC TTTAAAATCT GAACCTTCAA AGCTTGCTAT TAAATCATTT 3480  
 AGAATGTTTA CATTTACTAA GGTGTGCTGG GTCATGTAAA ATATTAGACA CTAATATTTT 3540  
 CATAGAAATT AGGCTGGAGA AAGAAGGAAG AAATGGTTTT CTTAAATACC TACAAAAAAG 3600  
 TTACTGTGGT ATCTATGAGT TATCATCTTA GCTGTGTTAA AAATGAATTT TTACTATGGC 3660



AGATATGGTA TGGATCGTAA AATTTTAAGC ACTAAAAATT TTTTCATAAC CTTTCATAAT 3720  
AAAGTTTAAT AATAGGTTTA TTAAGTGAAT TTCATTAGTT TTTTAAAAGT GTTTTTGGTT 3780  
TGTGTATATA TACATATACA AATACAACAT TTACAATAAA TAAATACTT GAAATTCTCA 3840  
AAAAAAAAAA AAAAAAAAAA AAAAA

Seq ID NO: 633 Protein sequence  
Protein Accession #: NP\_003807.1

1 11 21 31 41 51  
| | | | | |  
MGSGARFPSG TLRVRWLLLL GLVGPVLGAA RPFQQTSHL SSYEIITPWR LTRERREAPR 60  
PYSKQVSYVI QAEGKEHIIH LERNKDLLPE DFVVYTYNKE GTLITDHPNI QNHCHYRGYV 120  
EGVHNSSIAL SDCFLGRLL HLENASYGIE PLQNSSHFEH IIRMDVYK EPLKCGVSNK 180  
DIEKETAKDE EEEPPSMTQL LRRRRVLPQ TRYVELFIVV DKERYDMMGR NQTAVREEMI 240  
LLANYLDSMY IMLNIRIVLV GLEIWTNGNL INIVGGAGDV LGNFVQWREK FLITRRRHDS 300  
AQLVLKKGFG GTAGMAFVGT VCSRSHAGGI NVFGQITVET FASIVAHELG HNLGMNHDDG 360  
RDCSCGAKSC IMNSGASGSR NFSSCSAEDF EKLTNLKGGN CLLNIPKPDE AYSAPSCGNK 420  
LVDAGEECDG GTPKECELDG CCEGSTCKLK SFAECAYGDC CKDCRFLPGG TLRGKTSEC 480  
DVPEYCNSSS QFCQPDVFIQ NGYPCQNNKA YCNGMCQYY DAQCQVIFGS KAKAAPKDCF 540  
IEVNSKGRDF GNCGFSGNEY KKCAGTGNALC GKLCQENVQE IPVFGIVPAI IQTPSRGTC 600  
WGVDFQLGSD VPDPGMVNEG TKCGAGKICR NFQCVQASVL NYDCDVQKKC HGHGVCNSNK 660  
NCHCENGWAP PNCETKGYGG SVDSGPTYNE MNTALRDGLL VFFFLIVPLI VCAIFIFIKR 720  
DQLWRSYFRK KRSQTYESDG KNQANPSRQP GSVPRHVSVP TPPREVPIYA NRFAVPTYAA 780  
KQPQQFSPRP PPPQPKVSSQ GNLIPTAPAP APPLYSSLT

Seq ID NO: 634 DNA sequence  
Nucleic Acid Accession #: NM\_002091.1  
Coding sequence: 56..503

1 11 21 31 41 51  
| | | | | |  
AGTCTCTGCT CTTCCAGCC TCTCCGGCGC GCTCCAAGGG CTTCCCGTCG GGACCATGCG 60  
CGGCAGTGAG CTCCCGCTGG TCCTGCTGGC GCTGTCCTC TGCCTAGCGC CCCGGGGGCG 120  
AGCGGTCCCG CTGCTGCGG CGCGAGGGAC CGTGTGACC AAGATGTACC CGCGCGGCAA 180  
CCACTGGGCG GTGGGGCACT TAATGGGGAA AAAGAGCACA GGGGAGTCTT CTTCTGTTTC 240  
TGAGAGAGGG AGCCTGAAGC AGCAGCTGAG AGAGTACATC AGGTGGGAAG AAGCTGCAAG 300  
GAATTTGCTG GGTCTCATAG AAGCAAAGGA GAACAGAAAC CACCAGCCAC CTCACCCCAA 360  
GGCCTTGGGC AATCAGCAGC CTTCTGTTGA TTCAGAGGAT AGCAGCAACT TCAAAGATGT 420  
AGGTTCAAAA GGCAAAGTTG GTAGACTCTC TGCTCCAGGT TCTCAACGTG AAGGAAGGAA 480  
CCCCCAGCTG AACCAGCAAT GATAATGATG GCCTCTCTCA AAAGAGAAAA ACAAACCCC 540  
TAAGAGACTG AGTTCTGCAA GCATCAGTTC TACGGATCAT CAACAAGATT TCCTTGTCGA 600  
AAATATTTGA CTATTCTGTA TCTTTCATCC TTGACTAAAT TCGTGATTTT CAAGCAGCAT 660  
CTTCTGGTTT AAAGTTGTTT GCTGTGAACA ATGTGCGAAA AGAGTCTTCC AATTAATGCT 720  
TTTTTATATC TAGGCTACCT GTTGGTTAGA TTCAAGGCCC CGAGCTGTGA CCATTCACAA 780  
TAAAGCTTA AACACAT

Seq ID NO: 635 Protein sequence  
Protein Accession #: NP\_002082.1

1 11 21 31 41 51  
| | | | | |  
MRGSELPLVL LALVLCCLAPR GRAVPLPAGG GTVLTKMYPR GNHWAVGHLM GKKSTGESSS 60  
VSEKSLKQQ LREYIRWEEA ARNLLGLIEA KENRNHQPQ PKALGNQQPS WSEDSSNFK 120  
DVGSKGKVGR LSAPGSQREG RNPQLNQ

Seq ID NO: 636 DNA sequence  
Nucleic Acid Accession #: NM\_016522.1  
Coding sequence: 265..1299

1 11 21 31 41 51  
| | | | | |  
GCGGAAGCAG CGAGGAGGGA GCGCCCTTTG GCGGTCCTCC GTGGAACCGG TTTTCCGAGG 60  
CTGGCAAAAG CCGAGGCTGG ATTTGGGGGA GGAATATTAG ACTCGGAGGA GTCTGCGCGC 120  
TTTTCTCCTC CCCGCGCCTC CCGGTCGCGG CGGGTTCACC GCTCAGTCCC CGCGCTCGCT 180  
CCGCACCCCA CCACTTCCT GTGCTCGCCC GGGGGGCGTG TGCCGTGCGG CTGCCGAGT 240  
TCGGGGAAGT TGTGGCTGTC GAGAATGGGG GTCTGTGGGT ACCTGTTCTT GCCCTGGAAG 300  
TGCTCTGTTG TCGTGTCTCT CAGGCTGCTG TTCCTTGTAC CCACAGGAGT GCGGTCGCGC 360  
AGCGGAGATG CCACCTTCCC CAAAGCTATG GACAACGTGA CCGTCCGCGA GGGGGAGAGC 420  
GCCACCTCA GGTGCACTAT TGACAACCGG GTCACCCGGG TGGCTGGCT AAACCGCAGC 480  
ACCATCCTCT ATGCTGGGAA TGACAAGTGG TGCTGGATC CTGCGGTGGT CCTTCTGAGC 540  
AACACCCAAA CGCAGTACAG CATCGAGATC CACAACGTGG ATGTGTATGA CGAGGGCCCT 600  
TACACCTGCT CGGTGCAGAC AGACAACCAC CCAAGACCT CTAGGGTCCA CCTCATTTGT 660  
CAAGTATCTC CCAAAATTGT AGAGATTTCT TCAGATATCT CCATTAATGA AGGGAACAAT 720  
ATTAGCCTCA CCTGCATAGC AACTGGTAGA CCAGAGCCTA CGGTACTTGT GAGACACATC 780  
TCTCCCAAAG CGGTGGGCTT TGTGAGTGAA GACGAATACT TGGAAATTCA GGGCATCACC 840  
CGGGAACAGT CAGGGGACTA CGAGTGCAGT GCCTCCAATG ACGTGGCCGC GCGGCTGGTA 900  
CGGAGAGTAA AGGTACCCGT GAACTATCCA CCATACATTT CAGAAGCCAA GGTACAGGT 960  
GTCCCCGTGG GACAAAAGGG GACACTGCAG TGTGAAGCCT CAGCAGTCCC CTCAGCAGAA 1020  
TTCCAGTGGT ACAAGGATGA CAAAAGACTG ATTGAAGGAA AGAAAGGGGT GAAAGTGGAA 1080  
AACAGACCTT TCCTCTCAA ACTCATCTTC TTCAATGTCT CTGAACATGA CTATGGGAAC 1140  
TACACTTGCG TGGCTCCAA CAAGCTGGGC CACACCAATG CCAGCATCAT GCTATTTGGT 1200  
CCAGGCGCGG TCAGCGAGGT GAGCAACGGC ACGTCGAGGA GGGCAGGCTG CGTCTGGCTG 1260  
CTGCCTCTTC TGGTCTTGCA CCTGCTTCTC AAATTTTGAT GTGAGTGCCA CTTCCCCACC 1320  
CGGGAAGGGC TGCCGCCACC ACACCCACCA ACACAACAGC AATGGCAACA CCGACAGCAA 1380  
CCAATCAGAT ATATACAAAT GAAATTAGAA GAAACACAGC CTCATGGGAC AGAAATTTGA 1440  
GGGAGGGGAA CAAAGAATAC TTTGGGGGGA AAAGAGTTTT AAAAAAGAAA TTGAAAATTG 1500  
CCTTGACAGT ATTTAGGTAC AATGGAGTTT TCTTTTCCCA AACGGGAAGA ACACAGCACA 1560



CCGGCTTGG ACCCACTGCA AGCTGCATCG TGCAACCTCT TTGGTGCCAG TGTGGGCAAG 1620  
 GGCTCAGCCT CTCTGCCCAC AGACTGCCCC CACGTGGAAC ATTCTGGAGC TGGCCATCCC 1680  
 AAATTCATC AGTCCATAGA GACGAACAGA ATGAGACCTT CCGGCCCAAG CGTGGCGCTT 1740  
 CCGGCCCAAG CGTGGCGCTG CGGGCACTTT GGTAGACTGT GCCACCACGG CGTGTGTTGT 1800  
 GAAACGTGAA ATAAAAAGAG CAAAAAATAA AAAAAAAA

Seq ID NO: 637 Protein sequence  
 Protein Accession #: NP\_057606.1

1 11 21 31 41 51  
 | | | | | |  
 MGVCGYLFLP WKCLVVVSLR LLFLVPTGVP VRSGDATFPK AMDNVTVRQG ESATLRCTID 60  
 NRVTRVAWLN RSTILYAGND KWCLDPRVVL LSNTQTQYSI EIQNVDVYDE GPYTCSVQTD 120  
 NHPKTSRVHL IVQVSPKIVE ISSDISINEG NNISLTCIAT GRPEPTVTWR HISP KAVGFV 180  
 SEDEYLEIQG ITREQSGDYE CSASNDVAAP VVRVKVTVN YPPYISEAKG TGVPGVQKGT 240  
 LQCEASAVPS AEFQWKDDK RLIEGKKGVK VENRPFLSKL IFFNVSEHDY GNYTCVASNK 300  
 LGHTNASIML FGPVAVSEVS NGTSRRAGCV WLLPLLVLHL LLKF

Seq ID NO: 638 DNA sequence  
 Nucleic Acid Accession #: NM\_012261.1  
 Coding sequence: 203..1045

1 11 21 31 41 51  
 | | | | | |  
 GATTTGCTCT GCCAGCAGCT GTCGGTGCCG CGCTCGACAC CGAGTCCTAG CTAGGCGCTC 60  
 ACAGAAATACG CGCTCCCTCC CTCCCCCTTC TCTGTCCCCC GCCTCTCGCT CACCCCGGCC 120  
 CACTCCAGCG GCGACTTTGA GGGATTCCCT CTCTGGCGGC CTCTGCAGCA GCACAGCCGG 180  
 CCTCATTCGG GGCACGCGA GTATGGATCT CCAAGGAAGA GGGGTCCCCA GCATCGACAG 240  
 ACTTCGAGTT CTCCTGATGT TGTTCATAC AATGGCTCAA ATCATGGCAG AACAAGAAGT 300  
 GGAAAATCTC TCAGGCCTTT CCACTAACCC TGAAAAGAT ATATTTGTGG TCGGGGAAAA 360  
 TGGGACGACG TGTCTCATGG CAGAGTTTGC AGCCAAATTT ATTGTACCTT ATGATGTGTG 420  
 GGCCAGCAAC TACGTAGATC TGATCACAGA ACAGGCCGAT ATCGCATTGA CCCGGGGAGC 480  
 TGAGGTGAAG GGCCGCTGTG GCCACAGCCA GTCGGAGCTG CAAGTGTCTT GGGTGGATCG 540  
 CGCATATGCA CTCAAAATGC TCTTTGTAAA GGAAAGCCAC AACATGTCCA AGGGACCTGA 600  
 GGCGACTTGG AGGCTGAGCA AAGTGCAGTT TGTCTACGAC TCCTCGGAGA AAACCCACTT 660  
 CAAAGACGCA GTCAGTGCTG GGAAGCACAC AGCCAACTCG CACCACCTCT CTGCCTTGGT 720  
 CACCCCGCTT GGAAGTCCT ATGAGTGTCA AGCTCAACAA ACCATTTTAC TGGCCTCTAG 780  
 TGATCCGCGA AAGACGGTCA CCATGATCCT GTCTGCGGTC CACATCCAAC CTTTGTGACAT 840  
 TATCTCAGAT TTTGTCTTCA GTGAAGAGCA TAAATGCCCA GTGGATGAGC GGGAGCAACT 900  
 GGAAGAAACC TTGCCCCGTA TTTTGGGGCT CATCTTGGGC CTGTCATCA TGGTAACACT 960  
 CGCGATTTAC CACGTCCACC ACAAATGAC TGCCAACCAG GTGCAGATCC CTCGGGACAG 1020  
 ATCCCACTAT AAGCACATGG GCTAGAGGCC GTTAGGCAGG CACCCCTAT TCCTGCTCCC 1080  
 CCAACTGGAT CAGGTAGAAC AACAAAAGCA CTTTTCATC TTGTACACGA GATACACCAA 1140  
 CATAGCTACA ATCAAACAGG CCTGGGTATC TGAGGCTTGC TTGGCTTGTG TCCATGCTTA 1200  
 AACCCACGGA AGGGGGAGAC TCTTTCGGAT TTGTAGGGTG AAATGGCAAT TATTCTCTCC 1260  
 ATGCTGGGGG AGGGGGAGG AGGGTCTCAG ACAGCTTTCG TGCTCATGGT GGCTTGGCTT 1320  
 TGACTCTCCA AAGAGCAATA AATGCCACTT GGAGCTGTAT CTGGCCCCAA AGTTTAGGGA 1380  
 TTGAAAACAT GCTTCTTTGA GGAGGAAACC CTTTATAGTT CAGAAGAATA TGGGGTGTCT 1440  
 TGCTCCCTTG GACACAGCTG GCTTATCTTA TACAGTTGTC AATGCACACA GAATACAACC 1500  
 TCATGCTCCC TGCAGCAAGA CCCCTGAAAG TGATTCTATG TTCTGGCTGG CATCTGTCAT 1560  
 GTTTAGTGAT TGTCTTGGGA ATGTTTCACT GCTACCCGCA TCCAGCGACT GCAGCACCAG 1620  
 AAAACGACTA ATGTAACAT GCAGAGTTGT TTGGACTTCT TCCTGTGCCA GGTCCAAGTC 1680  
 GGGGGACCTG AAGAATCAAT CTGTGTGAGT CTGTTTTTCA AAATGAAATA AAACACACTA 1740  
 TTCTCTGGC

Seq ID NO: 639 Protein sequence  
 Protein Accession #: NP\_036393.1

1 11 21 31 41 51  
 | | | | | |  
 MDLQGRGVPS IDRLRVLLML FHTMAQIMAE QEVENLSGLS TNPEKDIFVV RENGTTCLMA 60  
 EFAAKFIVPY DVWASNYVDL ITEQADIALT RGAEVKGRCG HSQSELQVFW VDRAYALKML 120  
 FVKESHMSK GPEATWRLSK VQFVYDSSEK THFKDAVSAG KHTANSHHLS ALVTPAGKSY 180  
 ECQAQQTISL ASSDPQKTVT MILSAVHIQF FDIISDFVFS EEHKCPVDER EQLEETLPLI 240  
 LGLILGLVIM VTLAIYHVHH KMTANQVQIP RDRSQYKHM

Seq ID NO: 640 DNA sequence  
 Nucleic Acid Accession #: NM\_002993.1  
 Coding sequence: 64..408

1 11 21 31 41 51  
 | | | | | |  
 GGCACGAGCC AGTCTCCGCG CCTCCACCCA GCTCAGGAAC CCGCGAACCC TCTCTTGACC 60  
 ACTATGAGCC TCCCGTCCAG CCGCGCGGCC CGTGTCCCGG GTCCCTTCGGG CTCCTTGTGC 120  
 GCGCTGCTCG CGCTGCTGCT CCTGCTGACG CCGCCGGGGC CCCTCGCCAG CGCTGGTCCT 180  
 GTCTCTGCTG TGCTGACAGA GCTGCGTTGC ACTTGTTTAC GCGTTACGCT GAGAGTAAAC 240  
 CCCAAAACGA TTGGTAAACT GCAGGTGTTT CCGCGAGGCC CGCAGTGCTC CAAGGTGGAA 300  
 GTGGTAGCCT CCCTGAAGAA CCGGAAGCAA GTTTGTCTGG ACCCGGAAGC CCCTTTTCTA 360  
 AAGAAAGTCA TCCAGAAAAT TTTGGACAGT GGAAACAAGA AAAACTGAGT AACAAAAAAG 420  
 ACCATGCATC ATAAAATTGC CCAGTCTTCA GCGGAGCAGT TTTCTGGAGA TCCCTGGACC 480  
 CAGTAAGAAAT AAGAAGGAAG GGTGGTTTTT TTTCCATTTT CTACATGGAT TCCCTACTTT 540  
 GAAGAGTGTG GGGGAAAGCC TACGCTTCTC CTGAAGTTT ACAGCTCAGC TAATGAAGTA 600  
 CTAATATAGT ATTTCCACTA TTTACTGTTA TTTTACCTGA TAAGTTATTG AACCTTTTGG 660  
 CAATTGACCA TATTGTGAGC AAAGAATCAC TGGTTATTAG TCTTTCAATG AATATTGAAT 720  
 TGAAGATAAC TATTGTATTT CTATCATACA TTCCTTAAAG TCTTACCGAA AAGGCTGTGG 780  
 ATTTCTGATG GAAATAATGT TTTATTAGTG TGCTGTTGAG GGAGGTATCC TGTGTGTTCT 840  
 ACTCACTCTT CTCATAAAAT AGGAAATATT TTAGTTCTGT TTTCTTGGGG AATATGTTAC 900



TCTTTACCCCT AGGATGCTAT TTAAGTTGTA CTGTATTAGA ACACTGGGTG TGTACATACCG 960  
 TTATCTGTGC AGAATATATT TCCTTATTCA GAATTTCTAA AAATTTAAGT TCTGTAAGGG 1020  
 CTAATATATT CTCTTCCTAT GGTTTTAGAT GTTTGATGTC TTCTTAGTAT GGCATAATGT 1080  
 5 CATGATTTAC TCATTAACT TTTGATTTTGT ATGCTATTTT TTTACTATAG GATGACTATA 1140  
 ATTCTGGTCA CTAATATAC ACTTTAGATA GATGAAGAAG CCCAAAAACA GATAAATTCC 1200  
 TGATTGCTAA TTTACATAGA AATGTATTCT CTTGGTTT TAAATAAAG CAAAATTAAC 1260  
 AATGATCTGT GCTCTGCAA GTTTTGAAAA TATATTGAA CAATTTGAAT ATAAATTCAT 1320  
 CATTAGTCC TCAAAATATA TACAGCATTG CTAAGATTTT CAGATATCTA TTGTGGATCT 1380  
 10 TTTAAAGGTT TTGACCATTT TGTATGAGG AATTATACAT GTATCACATT CACTATATTA 1440  
 AAATTGCACT TTTATTTTTT CCTGTGTGTC ATGTTGGTTT TTGGTACTTG TATTGTCATT 1500  
 TGGAGAAACA ATAAAGATT TCTAAACCAA AAAAAA AAAAAA

Seq ID NO: 641 Protein sequence  
 Protein Accession #: NP\_002984.1

1 11 21 31 41 51  
 | | | | | |  
 MSLPSSRAAR VPGPSGSLCA LLALLLLLTP PGPLASAGPV SAVLTELRCT CLRVTLRVNP 60  
 KTIGKLQVFP AGPQCSKVEV VASLKNGKQV CLDPEAPFLK KVIQKILDSG NKKN

Seq ID NO: 642 DNA sequence  
 Nucleic Acid Accession #: NM\_013271.1  
 Coding sequence: 27..809

1 11 21 31 41 51  
 | | | | | |  
 TCCGGAGCCA GGCTCGCTGG GGCAGCATGG CGGGGTCGCC GCTGCTCTGG GGGCCGCGGG 60  
 CCGGGGGCGT CGGCCTTTTG GTGCTGCTGC TGCTCGGCCT GTTTCGGCCG CCCCCGCGC 120  
 30 TCTGCGCGCG GCCGGTAAAG GAACCCCGCG GCCTAAGCGC AGCGTCTCCG CCCTTGGCTG 180  
 AGACTGGCGC TCCTCGCCGC TTCCGGCGGT CAGTGCCCGG AGGTGAGGCG GCGGGGGCGG 240  
 TGCAGGAGCT GGCAGCGGCG CTGGCGCATC TGCTGGAGGC CGAACGTGAG GAGCGGGCGC 300  
 GGGCCGAGGC GCAGGAGGCT GAGGATCAGC AGGCGCGCGT CCTGGCGCAG CTGCTGCGCG 360  
 TCTGGGGCGC CCCCCGCAAC TCTGATCCGG CTCTGGGCCT GGACGACGAC CCCGACGCGC 420  
 CTGCAGCGCA GCTCGCTCGC GCTCTGCTCC CGGCCCGCCT TGACCTGCC GGCCTAGCAG 480  
 35 CCCAGCTTGT CCCCAGCGCC GTCCCCGCG CGGCGCTCCG ACCCCGGCCC CCGGTCTACG 540  
 ACGACGGCCC CGCGGGCCCC GATGCTGAGG AGGCAGGCGA CGAGACACCC GACGTGGACC 600  
 CCGAGCTGTT GAGGTACTTG CTGGGACGGA TTCTTGCGGG AAGCGCGGAC TCCGAGGGGG 660  
 TGGCAGCCCC GCGCCGCTTC CGCCGTGCGG CCGACCACGA TGTGGGCTCT GAGCTGCCCC 720  
 CTGAGGGCGT GCTGGGGGCG CTGCTGCGTG TGAAACGCCT AGAGACCCCG GCGCCCCAGG 780  
 40 TGCTGTCACG CCGCCTCTTG CCACCTGAG CACTGCCCGG ATCCCGTGCA CCCTGGGACC 840  
 CAGAAGTGCC CCCGCCATCC CGCCACCAGG ACTTCTCCCC GCCAGCACGT CCAGAGCAAC 900  
 TTACCCCGGC CAGCCAGCCC TCTCACCCGA GGATCCCTAC CCCCTGGCCC ACAATAACAT 960  
 GATCTGAGC

Seq ID NO: 643 Protein sequence  
 Protein Accession #: NP\_037403.1

1 11 21 31 41 51  
 | | | | | |  
 50 MAGSPLLWGP RAGGVGLLV LLLGLFRPPP ALCARPVKEP RGLSAASPPL AETGAPRRFR 60  
 RSVPRGEAAG AVQELARALA HLLAEERQER ARAEAQEAED QQARVLAQLL RVWGAPRNSD 120  
 PALGLDDDDP APAQLARAL LRARLDPAAL AAQLVPAPVP AAALRPRPPV YDDGPAGPDA 180  
 EEAGDETPDV DPELLRYLLG RILAGSADSE GVAAPRRLRR AADHDVGSSEL PPEGVLGALL 240  
 RVKRLETPAP QVPARRLLPP

Seq ID NO: 644 DNA sequence  
 Nucleic Acid Accession #: NM\_002214  
 Coding sequence: 681..2990

1 11 21 31 41 51  
 | | | | | |  
 CCCAGAGCCG CCTCCCCCTG TTGCTGGCAT CCCGAGCTTC CTCCCTTGCC AGCCAGGACG 60  
 CTGCCGACTT GTCTTTGCCG GCTGCTCCGC AGACGGGGCT GCAAAGCTGC AACTAATGGT 120  
 GTTGGCCTCC CTGCCACCT GTGGAAGCAA CTGCGCTGAT TGATGCGCCA CAGACTTTTT 180  
 65 TCCCTCGAC CTCGCCGCG TACCTCCCA CAGATCCAGC ATCACCCAGT GAATGTACAT 240  
 TAGGGTGGTT TCCCCCCCAG CTTGGGGCTT TGTTTGGGTT TGATTGTGTT TGGCTCTTCG 300  
 CTAAGCTGAT TTATGCAGCA GAAGCCCCAC CGGCTGGAGA GAAACAAAAG CTCTTTTCTT 360  
 TGTCCCGGAG CAGGCTGCGG AGCCCTTGCA GAGCCCTCTC TCCAGTCGCC GCCGGGCCCT 420  
 TGGCCGTCGA AGGAGGTGCT TCTCGCGGAG ACCGCGGGAC CCGCCGTGCC GAGCCGGGAG 480  
 70 GGCCGTAGGG GCCCTGAGAT GCCGAGCGGT GCCCGGGCCC GCTTACCTGC ACCGCTTGCT 540  
 CCGAGCCGCG GGGTCCGCTT GCTAGGCTGT CGGAACACGT CTTAGCGACA CTCGCCCCGCG 600  
 GGCCCCGAGG TCGCCCCGGA GGCCCGGCCC GCGTCCGGAA GGCAGCCAGG CGGCGGGCGC 660  
 GGGGCGGGCT GTTTTGCAAT ATGTGCGGCT CGGCCCTGGC TTTTTTTACC GCTGCATTTG 720  
 TCTGCCTGCA AAACGACCGG CGAGGTCCCG CCTCGTTCCT CTGGGCAGCC TGGGTGTTTT 780  
 75 CACTTGTCTT TGGACTGGGC CAAGGTGAAG ACAATAGATG TGCATCTTCA AATGCAGCAT 840  
 CCTGTGCCAG GTGCCTTGCG CTGGGTCCAG AATGTGGATG GTGTGTTCAA GAGGATTTCA 900  
 TTTCAAGTGG ATCAAGAAGT GAACGTTGTG ATATTGTTT CAATTTAATA AGCAAAGGCT 960  
 GCTCAGTTGA TTCAATAGAA TACCCATCTG TGCATGTTAT AATACCCACT GAAAATGAAA 1020  
 TTAATACCCA GGTGACACCA GGAGAAGTGT CTATCCAGCT GCGTCCAGGA GCCGAAGCTA 1080  
 80 ATTTTATGCT GAAAGTTCAT CCTCTGAAGA AATATCCTGT GGATCTTTAT TATCTTGTG 1140  
 ATGTCTCAGC ATCAATGCAC AATAATATAG AAAAATTAAA TTCCGTTGGA AACGATTTAT 1200  
 CTAGAAAAAT GGCATTTTTT TCCCGTGAAT TCTGCTTGG ATTTGGCTCA TACGTTGATA 1260  
 AAACAGTTTC ACCATACATT AGCATCCACC CCGAAAGGAT TCATAATCAA TGCAGTGAAT 1320  
 85 ACAATTTAGA CTGCATGCCT CCCCATGAG CCAATCATGT GCTGTCTTTG ACAGAGAACA 1380  
 TCACTGAGTT TGAGAAAGCA GTTCATAGAC AGAAGATCTC TGGAAACATA GATACACCAG 1440  
 AAGGAGGTTT TGACGCCATG CTTCAGGCAG CTGTCTGTGA AAGTCATATC GGATGGCGAA 1500  
 AAGAGGCTAA AAGATTGCTG CTGGTGATGA CAGATCAGAC GTCTCATCTC GCTCTTGATA 1560



5  
10  
15  
20  
25  
30  
35  
40

```

GCAAATTGGC AGGCATAGTG GTGCCCAATG ACGGAAACTG TCATCTGAAA AACAACGTCT 1620
ACGTCAAATC GACAACCATG GAACACCCCT CACTAGGCCA ACTTTCAGAG AAATTAATAG 1680
ACAACAACAT TAATGTCATC TTTGCAGTTC AAGGAAAACA ATTTTATTGG TATAAGGATC 1740
TTCTACCCCT CTTGCCAGGC ACCATTGCTG GTGAAATAGA ATCAAAGGCT GCAAACCTCA 1800
ATAATTTGGT AGTGGAAGCC TATCAGAAGC TCATTTTCTG AGTGAAAGTT CAGGTGGAAA 1860
ACCAGGTACA AGGCATCTAT TTTAACATTA CCGCCATCTG TCCAGATGGG TCCAGAAAGC 1920
CAGGCATGGA AGGATGCAGA AACGTGACGA GCAATGATGA AGTTCTTTTC AATGTAACAG 1980
TTACAATGAA AAAATGTGAT GTCACAGGAG GAAAAAATA TGCAATAATC AAACCTATTG 2040
GTTTTAATGA AACCGCTAAA ATTCATATAC ACAGAAACTG CAGCTGTCAG TGTGAGGACA 2100
ACAGAGGACC TAAAGGAAAG TGTGTAGATG AAACCTTTCT AGATTCCAAG TGTTCCTCAGT 2160
GTGATGAGAA TAAATGTCAT TTTGATGAAG ATCAGTTTTC TTCTGAGAGT TGCAAGTCAC 2220
ACAAGGATCA GCCTGTTTGC AGTGGTCGAG GAGTTTGTGT TTGTGGGAAA TGTTCATGTC 2280
ACAAAATTAA GCTTGGAAAA GTGTATGGAA AATACTGTGA AAAGGATGAC TTTTCTTGTC 2340
CATATCACCA TGGAAATCTG TGTGCTGGGC ATGGAGAGTG TGAAGCAGGC AGATGCCAAT 2400
GCTTCAGTGG CTGGGAAGGT GATCGATGCC AGTGCCCTTC AGCAGCAGCC CAGCACTGTG 2460
TCAATTCAAA GGGCCAAGTG TGCAGTGGAA GAGGCACGTG TGTGTGTGGA AGGTGTGAGT 2520
GCACCGATCC CAGGAGCATC GGCCGCTTCT GTGAACACTG CCCCACCTGT TATACAGCCT 2580
GCAAGGAAAA CTGGAATTGT ATGCAATGCC TTCACCTCA CAATTTGTCT CAGGCTATAC 2640
TTGATCAGTG CAAAACCTCA TGTGCTCTCA TGGAAACAAC GCATTATGTC GACCAAACCT 2700
CAGAATGTTT CTCCAGCCCA AGCTACTTGA GAATATTTT CATCATTTTC ATAGTTACAT 2760
TCTTGATTGG GTTGCTTAAA TTCCTGATG TTAGACAGGT GATACTACAA TGGAAATAGTA 2820
ATAAAATTAA GTCTCATCA GATTACAGAG TGTACGCTC AAAAAAGGAT AAGTTGATTC 2880
TGCAAAGTGT TTGCACAAGA GCAGTCACCT ACCGACGTGA GAAGCCTGAA GAAATAAAAA 2940
TGGATATCAG CAAATTAAT GCTCATGAAA CTTTCAGGTG CAACTTCTAA AAAAAGATT 3000
TTAAACACTT AATGGGAAAC TGGAAATTGT AATAATTGCT CCTAAAGATT ATAATTTTAA 3060
AAGTCACAGG AGGAGACAAA TTGCTCACGG TCATGCCAGT TGCTGGTTGT ACACTCGAAC 3120
GAAGACTGAC AAGTATCCTC ATCATGATG GACTCACATA GCTGCTGACT TTTTCAGAGA 3180
AAAATGTGTC TTACTACTGT TTGAGACTAG TGTCGTGTA GCACTTTACT GTAATATATA 3240
ACTTATTTAG ATCAGCATAG AATGTAGATC CTCTGAAGAG CACTGATTAC ACTTTACAGG 3300
TACCTGTTAT CCCTACGCTT CCCAGAGAGA ACAATGCTGT GAGAGAGTTT AGCATTGTGT 3360
CACTACAAGG GTACAGTAAT CCCTGCACTG GACATGTGAG GAAAAAATA ATCTGGCAAG 3420
TATATTCTAA GGTTGCCAAA CACTTCAACA GTTGGTGGTT GAATAGACAA GAACAGCTAG 3480
ATGAATAAAT GATTTCGTGT TCACTCTTTC AAGAGGTGAA CAGATACAAC CTTAATCTTA 3540
AAAGATTATT GCTTTTAAA GTGTGATGTT TTATGCAATG GTGTTTATGG TTTGCTTATT 3600
TTTGCAAGAT GGATACTAAT TCCAGCATTC TCTCCTCTTT GCCTTTATGT TTTGTTTCT 3660
TTTTTACAGG ATAAGTTTAT GTATGTCACA GATGACTGGA TTAATTAAGT GCTAAGTTAC 3720
TACTGCCATA AAAAATAAAT AATACAATGT CACTTTATCA GAATACTAGT TTTAAAAGCT 3780
GAATGTTAA

```

Seq ID NO: 645 Protein sequence  
Protein Accession #: NP\_002205

45  
50  
55  
60

```

1      11      21      31      41      51
|      |      |      |      |      |
MCGSALAFFT AAFVCLQNDR RGPASFLWAA WVFSVLVLGLG QGEDNRCASS NAASCARCLA 60
LGPECGWCVCV EDFISGGSR S ERCDIVSNLI SKGCSVDSIE YPSVHVIIPT ENEINTQVTP 120
GEVSIQLRPG AEANFMLKVH PLKKYPVDLY YLVDVSASMH NNIEKLNSVG NDLNRKMAFF 180
SRDFRLGFGS YVDKTVSPYI SIHPERIHNQ CSDYNLDCMP PHGYIHVLSL TENITEFEKA 240
VHRQKISGNI DTPEGGFDAM LQAAVCESHI GWRKEAKRL LVMTDQTS HL ALDSKLAGIV 300
VPNDGNCHLK NNVYVKSTTM EHPSLGQLSE KLIDNNINVI FAVQGKQFHW YKDLLPLLP 360
TIAGEIESKA ANLNNLVVEA YQKLISEVKV QVENQVQGIY FNITAICPDG SRKPGMEGCR 420
NVTSDNDELVF NVTVTMKKCD VTGGKNYAI KPIGFNETAK IHIHRNCSCQ CEDNRGPKGK 480
CVDETFLDSK CFQCDENKCH FDEDQFSSES CKSHKQDQVC SGRGVCVCGK CSCHKIKLGK 540
VYGKYCEKDD FSCPYHHGNL CAGHGECEAG RCQCFSGWEG DRCQCP SAAA QHCVNSKGQV 600
CSGRGTCVCG RCECTDPRSI GRFCEHCPTC YTACKENWNC MQCLHPHNLS QAILDQCKTS 660
CALMEQQHYV DQTSECFSSP SYLRIFFIIF IVTFLIGLLK VLIIRQVILQ WNSNRIKSSS 720
DYRVSASKKD KLILQSVCTR AVTYRREKPE EIKMDISKLN AHETFR CNF

```

Seq ID NO: 646 DNA sequence  
Nucleic Acid Accession #: NM\_003318.1  
Coding sequence: 1..2574

65  
70  
75  
80  
85

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGGAATCCG AGGATTTAAG TGGCAGAGAA TTGACAATTG ATTCCATAAT GAACAAAGTG 60
AGAGACATTA AAAATAAGTT TAAAAATGAA GACCTTACTG ATGAACTAAG CTTGAATAAA 120
ATTTCTGCTG ATACTACAGA TAACTCGGGA ACTGTTAACC AAATTATGAT GATGGCAAAC 180
AACCCAGAGG ACTGGTTGAG TTTGTTGCTC AAAC TAGAGA AAAACAGTGT TCCGCTAAGT 240
GATGCTCTTT TAAATAAAT GATTGGTCGT TACAGTCAAG CAATTGAAGC GCTTCCCCCA 300
GATAAATATG GCCAAATGA GAGTTTGTCT AGAATTC AAG TGAGATTGTC TGAATTA AAA 360
GCTATTCAAG AGCCAGATGA TGCAGTGAC TACTTTC AAG TGGCCAGAGC AAAC TGCAAG 420
AAATTTGCTT TTGTTTATAT ATCTTTTGCA CAATTTGAAC TGTCACAAGG TAATGTCAA 480
AAAAGTAAAC AACTTCTTCA AAAAGCTGTA GAACGTGGAG CAGTACCACT AGAAATGCTG 540
GAAATTGCCC TGCGGAATTT AAACCTCCAA AAAAGCAGC TGCTTTCAGA GGAGGAAAAG 600
AAGAATTTAT CAGCATCTAC GGTATTAAC TCCAGAGGAC AGACTACTAA AGCCAGGTTT 720
CATTTACAGA ATAGGAACAA CAGTTGTGAT TCCAGAGGAC AGACTACTAA AGCCAGGTTT 780
TTATATGGAG AGAACATGCC ACCACAAGAT GCAGAAATAG GTTACCGGAA TTCATTGAGA 840
CAAACTAACA AAATAAACA GTCATGCCCA TTTGGAAGAG TCCAGTTAA CCTTCTAAAT 900
AGCCCAGATT GTGATGTGAA GACAGATGAT TCAGTTGTAC CTTGTTTAT GAAAAGACAA 960
ACCTCTAGAT CAGAAATGCC AGATTGTGTT GTGCCTGGAT CTAACCAAG TGGAAATGAT 1020
TCCTGTGAAT TAAGAAATTT AAAGTCTGTT CAAAATAGTC ATTTCAAGGA ACCTCTGGTG 1080
TCAGATGAAA AGAGTTCTGA ACTTATTATT ACTGATTCAA TAACCTGAA GAATAAAACG 1140
GAATCAAGTC TTCTAGCTAA ATTAGAAGAA ACTAAAGAGT ATCAAGAAC AGAGGTTCCA 1200
GAGAGTAACC AGAAACAGTG GCAATCTAAG AGAAAGTCAG AGTGATTA TA CCAGAAATCCT 1260
GCTGCATCTT CAAATCACTG GCAGATTCCG GAGTTAGCCC GAAAAGTTAA TACAGAGCAG 1320
AAACATACCA CTTTGTAGCA ACCTGTCTTT TCAGTTTCAA AACAGTCACC ACCAATATCA 1380
ACATCTAAAT GGTGTGACCC AAAATCTATT TGTAAGACAC CAAGCAGCAA TACCTTGAT

```



	GATTACATGA	GCTGTTTTAG	AACTCCAGTT	GTAAAGAATG	ACTTTCCACC	TGCTTGTCAG	1440
	TTGTCAACAC	CTTATGGCCA	ACCTGCCTGT	TTCCAGCAGC	AACAGCATCA	AATACTTGCC	1500
	ACTCCACTTC	AAAATTACCA	GGTTTTAGCA	TCTTCTTCAG	CAAATGAATG	CATTTCCGGT	1560
5	AAAGGAAGAA	TTTATTCCAT	TTTAAAGCAG	ATAGGAAGTG	GAGGTTCAAG	CAAGGTATTT	1620
	CAGGTGTTAA	ATGAAAAGAA	ACAGATATAT	GCTATAAAAT	ATGTGAACTT	AGAAGAAGCA	1680
	GATAACCAAA	CTCTTGATAG	TTACCGGAAC	GAAATAGCTT	ATTTGAATAA	ACTACAACAA	1740
	CACAGTGATA	AGATCATCCG	ACTTTATGAT	TATGAAATCA	CGGACCAGTA	CATCTACATG	1800
	GTAATGGAGT	GTGGAAATAT	TGATCTTAAT	AGTTGGCTTA	AAAAGAAAAA	ATCCATTGAT	1860
10	CCATGGGAAC	GCAAGAGTTA	CTGGAAAAAT	ATGTTAGAGG	CAGTTCACAC	AATCCATCAA	1920
	CATGGCATTG	TTCACAGTGA	TCTTAAACCA	GCTAACTTTC	TGATAGTTGA	TGGAATGCTA	1980
	AAGCTAATTG	ATTTTGGGAT	TGCAAACCAA	ATGCAACCAG	ATACAACAAG	TGTTGTAAA	2040
	GATTCTCAGG	TTGGCACAGT	TAATTATATG	CCACCAGAA	CAATCAAAGA	TATGTCTTCC	2100
	TCCAGAGAGA	ATGGGAAATC	TAAGTCAAAG	ATAAGCCCCA	AAAGTGATGT	TTGGTCCTTA	2160
	GGATGTATTT	TGTACTATAT	GACTTACGGG	AAAACACCAT	TTCAGCAGAT	AATTAATCAG	2220
15	ATTTCTAAAT	TACATGCCAT	AATTGATCCT	AATCATGAAA	TTGAATTTCC	CGATATTCCA	2280
	GAGAAAGATC	TTCAAGATGT	GTTAAAGTGT	TGTTTAAAAA	GGGACCCAAA	ACAGAGGATA	2340
	TCCATTCCCTG	AGCTCCTGGC	TCATCCCTAT	GTTCAAATTC	AAACTCATCC	AGTTAACCAA	2400
	ATGGCCAAGG	GAACCACTGA	AGAAATGAAA	TATGTTCTGG	GCCAACCTGT	TGGTCTGAAT	2460
20	TCTCCTAACT	CCATTTTGAA	AGCTGCTAAA	ACTTTATATG	AACACTATAG	TGGTGGTGAA	2520
	AGTCATAATT	CTTCATCCTC	CAAGACTTTT	GAAAAAATAA	GGGGAAAAAA	ATGA	

Seq ID NO: 647 Protein sequence  
Protein Accession #: NP\_003309.1

25	1	11	21	31	41	51	
	MESEDLSGRE	LTIDSIMNKV	RDINKKFKNE	DLTDELSLNK	ISADTTDMSG	TVNQIMMMAN	60
	NPEDWLSLLL	KLEKNSVPLS	DALLNKLIGR	YSQAIEALPP	DKYGQNESFA	RIQVRFAELK	120
30	AIQEPDDARD	YFQMARANCK	KFAFVHISFA	QFELSQGNVK	KSKQLLQKAV	ERGAVPLEML	180
	EIALRNLNLQ	KKQLLSEEEK	KNLSASTVLT	AQESFSGSLG	HLQNRNNSCD	SRGQTTKARF	240
	LYGENMPPQD	AEIGYRNSLR	QTNKTKQSCP	FGRVPVNLN	SPDCDVKTDD	SVVPCFMKRQ	300
	TSRSECRDLV	VPGSKPSGND	SCELRNLKSV	QNSHFKEPLV	SDEKSSSELII	TDSITLKNKT	360
	ESSLLAKLEE	TKEYQEPFV	ESNQKQWQSK	RKSECINQNP	AASSNHWQIP	ELARKVNTEQ	420
	KHTTFEQPVF	SVSKQSPPI	TSKWFDPKSI	CKTPSSNTLD	DYMSCFRTPV	VKNDFPPACQ	480
35	LSTPYGQPAC	FQQQHQHILA	TPLQNLQVLA	SSSANECISV	KGRIYSILKQ	IGSGGSSKVF	540
	QVLNEKKQIY	AIKYVNLEEA	DNQTLDSYRN	EIAYLNKLQ	HSDKIIRLYD	YEITDQYIYM	600
	VMCEGNIDLN	SWLKKKSID	PWERKSYWKN	MLEAVHTIHQ	HGIVHSDLKP	ANFLIVDGML	660
	KLIDFGIANQ	MQPDTTSVVK	DSQVGTVNYM	PPEAIKDMSS	SRENGKSKSK	ISPKSDVWSL	720
40	GCILYYMTYG	KTPFQIINQ	ISKLHAIIDP	NHEIEFPDIP	EKDLQDVLKC	CLKRDPKQRI	780
	SIPELLAHPY	VQIQTHPVNQ	MAKGTTEEMK	YVLGQLVGLN	SPNSILKAAK	TLYEHYSGGE	840
	SHNSSSKTF	EKKRGKK					

Seq ID NO: 648 DNA sequence  
Nucleic Acid Accession #: NM\_015507  
Coding sequence: 241..1902

45	1	11	21	31	41	51	
50	CCGCAGAGGA	GCCTCGGCCA	GGCTAGCCAG	GGCGCCCCCA	GCCCCCTCCC	AGGCCGCGAG	60
	CGCCCCCTGCC	GCGGTGCCTG	GCCTCCCCTC	CCAGACTGCA	GGGACAGCAC	CCGGTAACTG	120
	CGAGTGGAGC	GGAGGACCCG	AGCGGCTGAG	GAGAGAGGAG	GCGGCGGCTT	AGCTGCTACG	180
	GGGTCCGGCC	GGCGCCCTCC	CGAGGGGGAG	TCAGGAGGAG	GAAGGAGGAC	CCGTGCGAGA	240
	ATGCCTCTGC	CCTGGAGCCT	TGCGCTCCCG	CTGCTGCTCT	CCTGGGTGGC	AGGTGGTTTC	300
55	GGGAACGCGG	CCAGTGCAAG	GCATCACGGG	TTGTTAGCAT	CGGCACGTCA	GCCTGGGGTC	360
	TGTCACTATG	GAACATAACT	GGCCTGCTGC	TACGGCTGGA	GAAGAAACAG	CAAGGGAGTC	420
	TGTGAAGCTA	CATGCGAACC	TGGATGTAAG	TTTGGTGAAG	GCGTGGGACC	AAACAAATGC	480
	AGATGCTTTC	CAGGATACAC	CGGGAATAAC	TGCAGTCAAG	ATGTGAATGA	GTGTGGAATG	540
	AAACCCCGGC	CATGCCAACA	CAGATGTGTG	AATACACACG	GAAGCTACAA	GTGCTTTTGC	600
60	CTCAGTGGCC	ACATGTCAT	GCCAGATGCT	ACGTGTGTGA	ACTCTAGGAC	ATGTGCCATG	660
	ATAAATGTCT	AGTACAGCTG	TGAAGACACA	GAAGAAGGGC	CACAGTGCCT	GTGTCCATCC	720
	TCAGGACTCC	GCCTGGCCCC	AAATGGAAGA	GACTGTCTAG	ATATTGATGA	ATGTGCCTCT	780
	GGTAAAGTCA	TCTGTCCCTA	CAATCGAAGA	TGTGTGAACA	CATTTGGAAG	CTACTACTGC	840
	AAATGTCACA	TTGGTTTCGA	ACTGCAATAT	ATCAGTGGAC	GATATGACTG	TATAGATATA	900
	AATGAATGTA	CTATGGATAG	CCATACGTGC	AGCCACCATG	CCAATTGCTT	CAATACCCAA	960
65	GGGTCTCTTA	AGTGTAATG	CAAGCAGGGA	TATAAAGGCA	ATGGACTTCG	GTGTTCTGCT	1020
	ATCCCTGAAA	ATTCTGTGAA	GGAAGTCTCT	AGAGCACCTG	GTACCATCAA	AGACAGAATC	1080
	AAGAAGTTGC	TTGCTCACA	AAACAGCATG	AAAAAGAAAG	CAAAAATTAA	AAATGTTACC	1140
	CCAGAACCCA	CCAGGACTCC	TACCCCTAAG	GTGAACTTGC	AGCCCTTCAA	CTATGAAGAG	1200
70	ATAGTTTCCA	GAGGCGGGAA	CTCTCATGGA	GGTAAAAAAG	GGAATGAAGA	GAAAATGAAA	1260
	GAGGGGCTTG	AGGATGAGAA	AAGAGAAGAG	AAAGCCCTGA	AGAATGACAT	AGAGGAGCGA	1320
	AGCCTGCGAG	GAGATGTGTT	TTTCCCTAAG	GTGAATGAAG	CAGGTGAATT	CGGCCTGATT	1380
	CTGGTCCAAA	GGAAAGCGCT	AACCTCCAAA	CTGGAACATA	AAGATTTAAA	TATCTCGGTT	1440
	GACTGCAGCT	TCAATCATGG	GATCTGTGAC	TGGAACACAG	ATAGAGAAGA	TGATTTTGAC	1500
75	TGGAATCCTG	CTGATCGAGA	TAATGCTATT	GGCTTCTATA	TGGCAGTTCC	GGCCTTGGCA	1560
	GGTCACAAGA	AAGACATTGG	CCGATTGAAA	CTTCTCTTAC	CTGACCTGCA	ACCCCAAAGC	1620
	AACTTCTGTT	TGCTCTTTGA	TTACCGGCTG	GCCGGAGACA	AAGTCGGGAA	ACTTCGAGTG	1680
	TTTGTGAAAA	ACAGTAACAA	TGCCCTGGCA	TGGGAGAAGA	CCACGAGTGA	GGATGAAAAG	1740
	TGGAAGACAG	GGAAATTTCA	GTTGTATCAA	GGAACATGAT	CTACCAAAAG	CATCATTTT	1800
80	GAAGCAGAAC	GTGGCAAGGG	CAAAACCGGC	GAAATCGCAG	TGGATGGCGT	CTTGCTTGTT	1860
	TCAGGCTTAT	GTCCAGATAG	CCTTTTATCT	GTGGATGACT	GAATGTTACT	ATCTTTATAT	1920
	TTGACTTTGT	ATGTCAGTTC	CCTGGTTTTT	TTGATATTGC	ATCATAGGAC	CTCTGGCATT	1980
	TTAGAATTAC	TAGCTGAAAA	ATTGTAATGT	ACCAACAGAA	ATATTATTGT	AAGATGCCTT	2040
	TCTTGTATAA	GATATGCCAA	TATTTGCTTT	AAATATCATA	TCACTGTATC	TTCTCAGTCA	2100
85	TTTCTGAATC	TTTCCACATT	ATATTGATAA	ATATGGAAAT	GTCAAGTTAT	CTCCCTCTCT	2160
	CAGTATATCT	GATTTGTATA	AGTAAGTTGA	TGAGCTTCTC	TCTACAACAT	TTCTAGAAAA	2220
	TAGAAAAAAA	AGCACAGAGA	AATGTTTAA	TGTTTGACTC	TTATGATACT	TCTTGGAAC	2280
	TATGACATCA	AAGATAGACT	TTTGCCTAAG	TGGCTTAGCT	GGGTCTTTCA	TAGCCAAACT	2340



TGTATATTTA AATTCTTTGT AATAATAATA TCCAAATCAT CAAAAAAAAA AAAAAAAA

Seq ID NO: 649 Protein sequence  
Protein Accession #: NP\_056322

1	11	21	31	41	51	
MPLPWSLALP	LLLSWVAGGF	GNAASARHHG	LLASARQPGV	CHYGTKLACC	YGWRRNSKGV	60
CEATCEPGCK	FGECVGPNNK	RCFPGYTGKT	CSQDVNECGM	KPRPCQHRCV	NTHGSYKCF	120
LSGHMLMPDA	TCVNSRTCAM	INCQYSCEDT	EEGPQCLCPS	SGLRLAPNGR	DCLDIDECAS	180
GKVICPYNRR	CVNTFGSYYC	KCHIGFELQY	ISGRYDCIDI	NECTMDSHTC	SHHANCNTQ	240
GSFKCKCKQG	YKGNLRCISA	IPENSVKEVL	RAPGTIKDRI	KKLLAHKNSM	KKKAKIKNVT	300
PEPTRTPPK	VNLQPFNYEE	IVSRGGNSHG	GKKGNEEKMK	EGLEDEKREE	KALKNDIEER	360
SLRGDVFFPK	VNEAGEFGLI	LVQRKALTSK	LEHKDLNISV	DCSFNHGICD	WKQDREDDFD	420
WNPADRDNAI	GFYMAVPALA	GHKKDIGRLK	LLLPDLQPQS	NFCLLFYDRL	AGDKVGKLRV	480
FVKNSNNALA	WEKTTSEDEK	WKTGKIQLYQ	GTDATKSIIF	EAERKGKGTG	EIAVDGVLLV	540
SGLCPDSLSS	VDD					

Seq ID NO: 650 DNA sequence  
Nucleic Acid Accession #: NM\_003506.1  
Coding sequence: 259..2379

1	11	21	31	41	51	
GCAGCTCCAG	TCCCGGACGC	AACCCCGGAG	CCGTCTCAGG	TCCCTGGGGG	GAACGGTGGG	60
TTAGACGGGG	ACGGGAAGGG	ACAGCGGCCT	TCGACCGCCC	CCCGAGTAAT	TGACCCAGGA	120
CTCATTTTCA	GGAAAGCCTG	AAAATGAGTA	AAATAGTGAA	ATGAGGAATT	TGAACATTTT	180
ATCTTTGGAT	GGGGATCTTC	TGAGGATGCA	AAGAGTGATT	CATCCAAGCC	ATGTGGTAAA	240
ATCAGGAATT	TGAAGAAAAT	GGAGATGTTT	ACATTTTGTG	TGACGTGTAT	TTTTCTACCC	300
CTCCTAAGAG	GGCACAGTCT	CTTCACCTGT	GAACCAATTA	CTGTTCCAG	ATGTATGAAA	360
ATGGCCTACA	ACATGACGTT	TTTCCCTAAT	CTGATGGGTC	ATTATGACCA	GAGTATTGCC	420
GCGGTGGAAA	TGGAGCATT	TCTTCTCTC	GCAAATCTGG	AATGTTTACC	AAACATTGAA	480
ACTTTCTCT	GCAAAAGCATT	TGTACCAACC	TGCATAGAAC	AAATTCATGT	GGTTCACCT	540
TGTCGTAAAC	TTTGTGAGAA	AGTATATTCT	GATTGCAAAA	AATTAATTGA	CACTTTTGGG	600
ATCCGATGGC	CTGAGGAGCT	TGAATGTGAC	AGATTACAAT	ACTGTGATGA	GACTGTTCCCT	660
GTAACTTTTG	ATCCACACAC	AGAATTTCTT	GGTCCCTCAG	AGAAAACAGA	ACAAGTCCAA	720
AGAGACATTG	GATTTTGGTG	TCCAAGGCAT	CTTAAGACTT	CTGGGGGACA	AGGATATAAG	780
TTTCTGGGAA	TTGACCAGTG	TGCGCCTCCA	TGCCCCAACA	TGTATTTTAA	AAGTGATGAG	840
CTAGAGTTTG	CAAAAAGTTT	TATTGGAACA	GTTCCTAATAT	TTTGTCTTTG	TGCAACTCTG	900
TTACATTTC	TTACTTTTTT	AATTGATGTT	AGAAGATTCA	GATACCCAGA	GAGACCAATT	960
ATATATTACT	CTGTCTGTTA	CAGCATTGTA	TCTCTTATGT	ACTTCATTGG	ATTTTTTGCTG	1020
GGCGATAGCA	CAGCCTGCAA	TAAGGCAGAT	GAGAAGCTAG	AACTTGGTGA	CACTGTTGTC	1080
CTAGGCTCTC	AAAATAAGGC	TTGCACCGTT	TTGTTTCATG	TTTTGTATTT	TTTCACAATG	1140
GCTGGCACTG	TGTGGTGGGT	GATCTTAC	ATTACTTGGT	TCTTAGCTGC	AGGAAGAAAA	1200
TGGAGTTGTG	AAGCCATCGA	GCAAAAAGCA	GTGTGTTTTC	ATGCTGTTGC	ATGGGGAACA	1260
CCAGGTTTCC	TGACTGTTAT	GCTTCTTGCT	CTGAACAAAG	TTGAAGGAGA	CAACATTAGT	1320
GGAGTTTGCT	TTGTTGGCCT	TTATGACCTG	GATGCTTCTC	GCTACTTTGT	ACTCTTGCCA	1380
CTGTGCCCTT	GTGTGTTTGT	TGGGCTCTCT	CTTCTTTTAG	CTGGCATTAT	TTCTTTAAAT	1440
CATGTTTCGAC	AAGTCATACA	ACATGATGGC	CGGAACCAAG	AAAAACTAAA	GAAATTTATG	1500
ATTGCAATTG	GAGTCTTCAG	CGGCTTGTAT	CTTGTGCCAT	TAGTGACACT	TCTCGGATGT	1560
TACGTCTATG	AGCAAGTGAA	CAGGATTACC	TGGGAGATAA	CTTGGGTCTC	TGATCATTGT	1620
CGTCAGTACC	ATATCCCATG	TCCTTATCAG	GCAAAAGCAA	AAGCTCGACC	AGAATTGGCT	1680
TTATTTATGA	TAAAATACCT	GATGACATTA	ATTGTTGGCA	TCTCTGCTGT	CTTCTGGGTT	1740
GGAAGCAAAA	AGACATGCAC	AGAATGGGCT	GGGTTTTTTA	AACGAAATCG	CAAGAGAGAT	1800
CCAATCAGTG	AAAGTCGAAG	AGTACTACAG	GAATCATGTG	AGTTTTTCTT	AAAGCACAAAT	1860
TCTAAAGTTA	AACACAAAAA	GAAGCACTAT	AAACCAAGTT	CACACAAGCT	GAAGGTCATT	1920
TCCAAATCCA	TGGGAACCGA	CACAGGAGCT	ACAGCAAAATC	ATGGCACTTC	TGCAGTAGCA	1980
ATTACTAGCC	ATGATTACCT	AGGACAAGAA	ACTTTGACAG	AAATCCAAAC	CTCACCAGAA	2040
ACATCAATGA	GAGAGGTGAA	AGCGGACGGA	GCTAGCACCC	CCAGGTAAAG	AGAACAGGAC	2100
TGTGGTGAAC	CTGCCCTGCC	AGCAGCATCC	ATCTCCAGAC	TCTCTGGGGA	ACAGGTCGAC	2160
GGGAAGGGCC	AGGCAGGCAG	TGTATCTGAA	AGTGCGCGGA	GTGAAGGAAG	GATTAGTCCA	2220
AAGAGTGATA	TACTGACAC	TGGCCTGGCA	CAGAGCAACA	ATTTGCAGGT	CCCCAGTTCT	2280
TCAGAACCAA	GCAGCCTCAA	AGGTTCCACA	TCTCTGCTTG	TTACCCAGT	TTCAGGAGTG	2340
AGAAAAGAGC	AGGGAGGTGG	TTGTCTATCA	GATACCTGAA	GAACATTTTC	TCTCGTTACT	2400
CAGAAGCAAA	TTTGTGTTAC	ACTGGAAGTG	ACCTATGCAC	TGTTTTGTAA	GAATCACTGT	2460
TACGTTCTTC	TTTTGCACTT	AAAGTTGCAT	TGCCTACTGT	TATACTGGAA	AAAATAGAGT	2520
TCAAGAATAA	TATGACTCAT	TTACACAAAA	GGTTAATGAC	AACAATATAC	CTGAAAACAG	2580
AAATGTGCAG	GTTAATAATA	TTTTTTTAAAT	AGTGTGGGAG	GACAGAGTTA	GAGGAATCTT	2640
CCTTTTCTAT	TTATGAAGAT	TCTACTCTTG	GTAAGAGTAT	TTAAGATGT	ACTATGCTAT	2700
TTTACCTTTT	TGATATAAAA	TCAAGATATT	TCTTTGCTGA	AGTATTTAAA	TCTTATCCTT	2760
GTATCTTTTT	ATACATATTT	GAAAATAAGC	TTATATGTAT	TTGAACCTTT	TTGAAATCCT	2820
ATTCAAGTAT	TTTTATCATG	CTATTGTGAT	ATTTTAGCAC	TTTGGTAGCT	TTTACACTGA	2880
ATTTCTAAGA	AAATTGTAAA	ATAGTCTTCT	TTTATACTGT	AAAAAAGAT	ATACCAAAAA	2940
GTCTTATAAT	AGGAATTTAA	CTTTAAAAAC	CCACTTATTG	ATACCTTACC	ATCTAAAATG	3000
TGTGATTTTT	ATAGTCTCGT	TTTAGGAATT	TCACAGATCT	AAATTATGTA	ACTGAAATAA	3060
GGTGCTTACT	CAAAGAGTGT	CCACTATTGA	TTGTATTATG	CTGCTCACTG	ATCCTTCTGC	3120
ATATTTAAAA	TAAAATGTCC	TAAAGGGTTA	GTAGACAAAA	TGTTAGTCTT	TTGTATATTA	3180
GGCCAAGTGC	AATTGACTTC	CCTTTTTTAA	TGTTTCATGA	CCACCCATTG	ATTGTATTAT	3240
AACCACTTAC	AGTTGCTTAT	ATTTTGTGTT	TTAACCTTTG	TTTCTTAACA	TTTAGAATAT	3300
TACATTTTGT	ATTATACAGT	ACCTTTCTCA	GACATTTTGT	AG		

Seq ID NO: 651 Protein sequence  
Protein Accession #: NP\_003497.1

1	11	21	31	41	51



MEMFTFL LTC IFLPLLRGHS LFTCEPITVP RCMK MAYNMT FFPNLMGHYD QSIAAVEMEH 60  
 FLPLANLECS PNIETFLCKA FVPTCIEQIH VVPPCRKLCE KVYSDCKKLI DTFGIRWPDEE 120  
 LECDRLQYCD ETVPVTFDPH TEFLGPQKKT EQVQRDIGFW CPRHLKTS GG QGYKFLGIDQ 180  
 5 CAPP CPNMYF KSDELEFAKS FIGTVSIFCL CATLFTFLTF LIDVRRFRYP ERPIIYYSVC 240  
 YSIVSLMYFI GFL LGDSTAC NKADEKLELG DTVVLG SQNK ACTVLFMLLY FFTMAGTVWW 300  
 VILTTIWFLA AGRKWSCEAI EQKAVWFHAV AWGTPGF LTV MLLALNKVEG DNISGVC FVG 360  
 LYDL DASRYF VLLPLCLCVF VGLSLLLAGI ISLNHVRQVI QHDGRNQEKL KKF MIRIGVF 420  
 SGLYL VPLVT LLGCYVYEQV NRITWEITWV SDHCRQYHIP CPYQAKAKAR PELALFMIKY 480  
 10 LMTLIVGISA VFWVGSKKT C TEWAGFFKR N RKRDPISESR RVLQESCEFF LKHNSKVHKH 540  
 KKHYKPSSH K LKVISKSMGT STGATANHGT SAVAITSHDY LGQETL TEIQ TSPETSMREV 600  
 KADGASTPRL REQDCGEPAS PAASISRLSG EQVDGKGQAG SVSESARSEG RISPKSDITD 660  
 TGLAQSNNLQ VPSSSEPSSL KGSTSLLVHP VSGVRKEQGG GCHSDT

Seq ID NO: 652 DNA sequence  
 Nucleic Acid Accession #: NM\_014791.1  
 Coding sequence: 171..2126

1 11 21 31 41 51  
 | | | | |  
 20 TTGGCGGGCG GAAGCGGCCA CAACCCGGCG ATCGAAAAGA TTCTTAGGAA CGCCGTACCA 60  
 GCCGCTCTC TCAGGACAGC AGGCCCTGT CTTCTGTG GCGCCGCTC AGCCGTGCC 120  
 TCCGCCCTC AGGTTCTTT TCTAATTCCA AATAAACTTG CAAGAGGACT ATGAAAGATT 180  
 ATGATGAAC TCTCAATAT TATGAATTAC ATGAACTAT TGGGACAGGT GGCTTTGCAA 240  
 25 AGGTCAAAC TGCCTGCCAT ATCCTTACTG GAGAGATGGT AGCTATAAAA ATCATGGATA 300  
 AAAACACACT AGGGAGTGAT TTGCCCCGGA TCAAAACGGA GATTGAGGCC TTGAAGAACC 360  
 TGAGACATCA GCATATATGT CAACTCTACC ATGTGCTAGA GACAGCCAAC AAAATATTCA 420  
 TGGTTCTTGA GTACTGCCCT GGAGGAGAGC TGTGTTGACTA TATAATTTCC CAGGATCGCC 480  
 TGT CAGAAGA GGAGACCCGG GTTGTCTTCC GT CAGATAGT ATCTGCTGTT GCTTATGTGC 540  
 ACAGCCAGGG CTATGCTCAC AGGGACCTCA AGCCAGAAAA TTTGCTGTTT GATGAATATC 600  
 30 ATAAATTA GCTGATTGAC TTTGGTCTCT GTGCAAAACC CAAGGGTAAC AAGGATTACC 660  
 ATCTACAGAC ATGCTGTGGG AGTCTGGCTT ATGCAGCACC TGAGTTAATA CAAGGCAAAT 720  
 CATATCTTGG ATCAGAGGCA GATGTTTGGA GCATGGGCAT ACTGTTATAT GTTCTTATGT 780  
 GTGGATTCT ACCATTTGAT GATGATACT TAATGGCTTT ATACAAGAAG ATTATGAGAG 840  
 35 GAAAATATGA TGTTCCTAAG TGGCTCTCTC CCAGTAGCAT TCTGCTTCTT CAACAAATGC 900  
 TGCAGTGGA CCCAAAGAAA CGGATTTCTA TGAAAAATCT ATTGAACCAT CCCTGGATCA 960  
 TGCAAGATTA CAACTATCCT GTTGAGTGGC AAAGCAAGAA TCCTTTTATT CACCTCGATG 1020  
 ATGATTGCGT AACAGAACTT TCTGTACATC ACAGAAACAA CAGGCAAAACA ATGGAGGATT 1080  
 TAATTTCACT GTGGCAGTAT GATCACCTCA CGGCTACCTA TCTTCTGCTT CTAGCCAAGA 1140  
 AGGCTCGGGG AAAACCAAGT CGTTTAAGGC TTTCTTCTTT CTCTGTGGA CAAGCCAGTG 1200  
 40 CTACCCCATT CACAGACATC AAGTCAAATA ATTGAGTCT GGAAGATGTG ACCGCAAGTG 1260  
 ATAAAAATTA TGTGGCGGGA TTAATAGACT ATGATTGGTG TGAAGATGAT TTATCAACAG 1320  
 GTGCTGCTAC TCCCCGAACA TCACAGTTTA CCAAGTACTG GACAGAATCA AATGGGGTGG 1380  
 AATCTAAATC ATTAACCTCA GCCTTATGCA GAACACCTGC AAATAAATTA AAGAACAAAG 1440  
 45 AAAATGTATA TACTCCTAAG TCTGCTGTAA AGAATGAAGA GTACTTTATG TTTCTGAGC 1500  
 CAAAGACTCC AGTTAATAAG AACCAGCATA AGAGAGAAAT ACTCACTACG CCAAATCGTT 1560  
 AACTACACC CTCAAAGCT AGAAACCACT GCCTGAAAGA AACTCCAATT AAAATACCAG 1620  
 TAAATTC AACAGACAGC AAGTTAATGA CAGGTTCAT TAGCCCTGAG AGGCGGTGCC 1680  
 GCTCAGTGGA ATTGGATCTC AACCAAGCAC ATATGGAGGA GACTCCAAA AGAAAGGGAG 1740  
 CCAAAGTGT TGGGAGCCTT GAAAGGGGGT TGGATAAGGT TATCACTGTG CTCACCAGGA 1800  
 50 GCAAAGGAA GGGTTCTGCC AGAGACGGGC CCAGAAAGCT AAAGCTTCAC TATAATGTGA 1860  
 CTACAACTAG ATTAGTGAAT CCAGATCAAC TGTGTAATGA AATAATGTCT ATTCTTCCAA 1920  
 AGAAGCATGT TGACTTTGTA CAAAAGGGT ATACTGAA GTGTCAAACA CAGTCAGATT 1980  
 TTGGGAAAGT GACAATTCAA TTTGAATTAG AAGTGTGCCA GCTTCAAAA CCCGATGTGG 2040  
 55 TGGGTATCAG GAGGCAGCGG CTTAAGGGCG ATGCCTGGGT TTACAAAAGA TTAGTGGAG 2100  
 ACATCCTATC TAGCTGCAAG GTATAATTGA TGGATTCTTC CATCTGCCG GATGAGTGTG 2160  
 GGTGTGATAC AGCCTACATA AAGACTGTTA TGATCGCTTT GATTTTAAAG TTCATTGGAA 2220  
 CTACCAACTT GTTTCTAAAG AGCTATCTTA AGACCAATAT CTCTTTGTTT TTAACAAAA 2280  
 GATATTATTT TGTGTATGAA TCTAAATCAA GCCCATCTGT CATTATGTTA CTGTCTTTT 2340  
 60 TAATCATGTG GTTTGTATA TTAATAATTG TTGACTTTCT TAGATTCACT TCCATATGTG 2400  
 AATGTAAGCT CTTAACTATG TCTCTTTGTA ATGTGTAATT TCTTTCTGAA ATAAACCAT 2460  
 TTGTGAATAT

Seq ID NO: 653 Protein sequence  
 Protein Accession #: NP\_055606.1

1 11 21 31 41 51  
 | | | | |  
 65 MKDYDELLKY YELHETIGTG GFAKVKLACH ILTGEMVAIK IMDKNTLGSD LPRIKTEIEA 60  
 LKNLRHQHIC QLYHVLETAN KIFMVLEYCP GGELFDYIIS QDR LSEETR VVFRQIVSAV 120  
 70 AYVHSQGYAH RDLKPENLLF DEYHKLKLID FGLCAKPKGN KDYHLQTCG SLAYAAP ELI 180  
 QGKSYLGSEA DVWSMGILLY VLMCGFLFPD DDNVMALYKK IMRGKYDVPK WLS PSSL LLL 240  
 QQMLQVDPKK RISMKNLLNH PWIMQDYNYP VEWQSKNPF I HLDDDCVTEL SVHHRNNRQT 300  
 MEDLISLWQY DHLTATYLLL LAKKARGKPV RLRLSSFSCG QASATPFTDI KSNNWSLEDV 360  
 75 TASDKNYVAG LIDYDWCEDD LSTGAATPRT SQFTKYWTES NGVESKSLTP ALCRTPANKL 420  
 KNKENVYTPK SAVKNEEYFM FPEPKTPVNK NQHKREILT PNR YTTPSKA RNQCLKETPI 480  
 KIPVNSTGTD KLMTGVISPE RRCRSVELDL NQAHMEETPK RKGAKVFGSL ERGLDKVITV 540  
 LTRSKRK GSA RDGPRRLKLH YNVTTT RLVN PDQLNEIMS ILPKKHVDFV QKGYTLKCQT 600  
 QSDFGKVTM Q FELEVCLQK PDVVGIR RQR LKGDWVYKR LVEDILSSCK V

Seq ID NO: 654 DNA sequence  
 Nucleic Acid Accession #: NM\_000582  
 Coding sequence: 88..990

1 11 21 31 41 51  
 | | | | |  
 85 GCAGAGCACA GCATCGTCGG GACCAGACTC GTCTCAGGCC AGTTGCAGCC TTCTCAGCCA 60  
 AACGCCGACC AAGGAAAAC TACTACCATG AGAATTGCAG TGATTGCTT TTGCCTCCTA 120



5  
10  
15  
20  
25

```

GGCATCACCT GTGCCATACC AGTTAAACAG GCTGATTCTG GAAGTTCTGA GGAAAAGCAG 180
CTTTACAACA AATACCCAGA TGCTGTGGCC ACATGGCTAA ACCCTGACCC ATCTCAGAAG 240
CAGAATCTCC TAGCCCCACA GACCCCTCCA AGTAAGTCCA ACGAAAGCCA TGACCACATG 300
GATGATATGG ATGATGAAGA TGATGATGAC CATGTGGACA GCCAGGACTC CATTGACTCG 360
AACGACTCTG ATGATGTAGA TGACACTGAT GATTCTCACC AGTCTGATGA GTCTCACCAT 420
TCTGATGAAT CTGATGAACT GGTCACTGAT TTCCCCACGG ACCTGCCAGC AACCAGAGTT 480
TTCACTCCAG TTGTCCCCAC AGTAGACACA TATGATGGCC GAGGTGATAG TGTGGTTTAT 540
GGACTGAGGT CAAAATCTAA GAAGTTTCGC AGACCTGACA TCCAGTACCC TGATGCTACA 600
GACGAGGACA TCACCTCACA CATGGAAAGC GAGGAGTTGA ATGGTGCATA CAAGGCCATC 660
CCCGTTGCCC AGGACCTGAA CGCGCCTTCT GATTGGGACA GCCGTGGGAA GGACAGTTAT 720
GAAACGAGTC AGCTGGATGA CCAGAGTGCT GAAACCCACA GCCACAAGCA GTCCAGATTA 780
TATAAGCGGA AAGCCAATGA TGAGAGCAAT GAGCATTCGG ATGTGATTGA TAGTCAGGAA 840
CTTTCCAAAG TCAGCCGTGA ATTCCACAGC CATGAATTTT ACAGCCATGA AGATATGCTG 900
GTTGTAGACC CCAAAAGTAA GGAAGAAGAT AAACACCTGA AATTTTCGTAT TTCTCATGAA 960
TTAGATAGTG CATCTTCTGA GGTCAATTAA AAGGAGAAAA AATACAATTT CTCACTTTGC 1020
ATTTAGTCAA AAGAAAAAAT GCTTTATAGC AAAATGAAAG AGAACATGAA ATGCTTCTTT 1080
CTCAGTTTAT TGGTTGAATG TGTATCTATT TGAGTCTGGA AATAACTAAT GTGTTTGATA 1140
ATTAGTTTAG TTTGTGGCTT CATGGAAACT CCCTGTAAAC TAAAAGCTTC AGGGTTATGT 1200
CTATGTTTAT TCTATAGAAG AAATGCAAAC TATCACTGTA TTTTAATATT TGTATTCTC 1260
TCATGAATAG AAATTTATGT AGAAGCAAAC AAAATACTTT TACCCACTTA AAAAGAGAAT 1320
ATAACATTTT ATGTCACTAT AATCTTTTGT TTTTAAAGTT AGTGATATAT TTGTTGTGAT 1380
TATCTTTTGT TGGTGTGAAT AAATCTTTTA TCTTGAATGT AATAAGAATT TGGTGGTGTC 1440
AATTGCTTAT TTGTTTTCCC ACGGTTGTCC AGCAATTAAT AAAACATAAC CTTTTTTACT 1500
GCCTAAAAAA AAAAAAAAAA AAAA

```

Seq ID NO: 655 Protein sequence  
Protein Accession #: NP\_000573

30  
35

```

1 11 21 31 41 51
| | | | |
MRIAVICFCL LGITCAIPVK QADSGSSEEEK QLYNKYPDAV ATWLNPDPSQ KQNLLAPQTL 60
PSKSNESH DH MDDMDEDD DHVDSQDSID SNDSDDVDDT DDSHQSDSH HSDESDELVT 120
DFPTDL PATE VFTPVVPTVD TYDGRGDSV YGLRSKSKKF RRPDIQYPDA TDEDITSHME 180
SEELNGAYKA IPVAQDLNAP SDWDSRGKDS YETSQLDDQS AETHSHKQSR LYKRKANDES 240
NEHSDVIDSQ ELSKVSREFH SHEFHSHEM LVVDPKSKEE DKHLKFRISH ELDSASSEVN

```

Seq ID NO: 656 DNA sequence  
Nucleic Acid Accession #: NM\_003108.1  
Coding sequence: 76..1401

40  
45  
50  
55  
60  
65  
70  
75  
80

```

1 11 21 31 41 51
| | | | |
GGGGTGGGAG GGGGAGGGGG ACCTCCGCAC GAGACCCAGC GGCCCGGGTT GGAGCGTCCA 60
GCCCTGCAAC GGATCATGGT GCAGCAGGCG GAGAGCTTGG AAGCGGAGAG CAACCTGCCC 120
CGGGAGGCGC TGGACACGGA GGAGGCGGAA TTCATGGCTT GCAGCCCGGT GGCCCTGGAC 180
GAGAGCGACC CAGACTGGTG CAAGACGGCG TCGGGCCACA TCAAGCGGCC GATGAACGCG 240
TTCATGGTAT GGTCCAAGAT CGAACGCAGG AAGATCATGG AGCAGTCTCC GGACATGCAC 300
AACGCCGAGA TCTCCAAGAG GCTGGGCAAG CGCTGGAAAA TGCTGAAGGA CAGCGAGAAG 360
ATCCCGTTCA TCCGGGAGGC GGAGCGGCTG CGGCTCAAGC ACATGGCCGA CTACCCCGAC 420
TACAAGTACC GGCCCCGGAA AAAGCCCAAA ATGGACCCCT CGGCCAAGCC CAGCGCCAGC 480
CAGAGCCGAG AGAAGAGCGC GGCCGGCGGC GCGCGGCGGC AGGCGCGGGC 540
GGTGCCAAAGA CCTCCAAGGG CTCCAGCAAG AAATGCGGCA AGCTCAAGGC CCCC GCGGGC 600
GCGGGCGCCA AGGCGGGCGC GGGCAAGGCG GCCCAGTCCG GGGACTACGG GGGCGCGGGC 660
GACGACTACG TGCTGGGCAG CCTGCGCGTG AGCGGCTCGG GCGGCGGCGG CGCGGGCAAG 720
ACGGTCAAGT GCGTGTCTTCT GGATGAGGAC GACGACGACG ACGACGACGA CGACGAGCTG 780
CAGCTGCAGA TCAAAACAGGA GCCGGACGAG GAGGACGAGG AACCACCGCA CCAGCAGCTC 840
CTGCAGCCGC CGGGGCAGCA GCCGTCGAG CTGCTGAGAC GCTACAACGT CGCCAAAGTG 900
CCCGCCAGCC CTACGCTGAG CAGCTCGGCG GAGTCCCCCG AGGGAGCGAG CCTCTACGAC 960
GAGGTGCGGG CCGGCGCGAC CTCGGGCGCC GGGGGCGGCA GCCGCTCTA CTACAGCTTC 1020
AAGAACATCA CCAAGCAGCA CCCGCCGCCG CTCGCGCAGC CCGCGCTGTC GCCCGCGTCC 1080
TCGCGCTCGG TGTCCACCTC CTCGTCCAGC AGCAGCGGCA GCAGCAGCGG CAGCAGCGGC 1140
GAGGACGCGC ACGACCTGAT GTTCGACCTG AGCTTGAATT TCTCTCAAAG CGCGCACAGC 1200
GCCAGCGAGC AGCAGCTGGG GGGCGGCGGC GCGGCCGGGA ACCTGTCCCT GTCGCTGGTG 1260
GATAAGGATT TGGATTCTGT CAGCGAGGCG AGCCTGGGCT CCCACTTCGA GTTCCCCGAC 1320
TACTGCACGC CGGAGCTGAG CGAGATGATC GCGGGGGACT GGCTGGAGGC GAACTTCTCC 1380
GACCTGGTGT TCACATATTG AAAGGCGCCC GCTGCTCGCT CTTTCTCTCG GAGGGTGCAG 1440
AGCTGGGTTC CTGGGAGGA AGTTGTAGTG GTGATGATGA TGATGATGAT AATGATGATG 1500
ATGATGGTGG TGTGATGGT GGCGGTGGTA GGGTGGAGGG GAGAGAAGAA GATGCTGATG 1560
ATATTGATAA GATGTCGTGA CGCAAAGAAA TTGGAAACA TGATGAAAT TTTGGTGGAG 1620
TTAAAGTGAA ATGAGTAGTT TTTAAACATT TTTCTGTCC TTTTTTGTG CCCCCCTCC 1680
TCCTTTATCG TGTCTCAAGG TAGTTGCATA CCTAGTCTGG AGTTGTGATT ATTTTCCCAA 1740
AAAATGTGTT TTTGTAATTA CTATTTCTTT TTCTGAAAT TCGTGATTGC AACAAAGGCA 1800
GAGGGGGCGG CGCGGCGGAG GGGAGGTAGG ACCCGCTCCG GAAGGCGCTG TTTGAAGCTT 1860
GTCGGTCTTT GAAGTCTGGA AGACGTCTGC AGAGGACCCCT TTTGGCAGCA CAACTGTTAC 1920
TCTAGGGAGT TGGTGGAGAT ATTTTTTTTT CTTAAGAGAA CTTAAAGAAC TGGTGATTTT 1980
TTTTTAACAA AAAAAGGG

```

Seq ID NO: 657 Protein sequence  
Protein Accession #: NP\_003099.1

80  
85

```

1 11 21 31 41 51
| | | | |
MVQQAESLEA ESNLPREALD TEEGEFMACS PVALDESDPD WCKTASGHIK RPMNAFMVWS 60
KIERRKIMEQ SPDMHNAEIS KRLGKRWKML KDSEKIPFIR EAERLRLKHM ADYDPYKYRP 120
RKKPKMDPSA KPSASQSPEK SAAGGGGSGA GGGAGGAKTS KGSSKKCKGL KAPAAAGAKA 180
GAGKAAQSGD YGGAGDDYVL GSLRVSGSGG GGAGKTVKCV FLDEDDDDDD DDDELQLQIK 240

```



QEPDEEDEEP PHQQLLQPPG QQPSQLLRRY NVAKVPASPT LSSSAESPEG ASLYDEVVRAG 300  
 ATSGAGGGSR LYYSFKNITK QHPPPLAQPA LSPASSRSVS TSSSSSSGSS SGSSGEDADD 360  
 LMFDSLNLNF QSAHSASEQQ LGGGAAAGNL SLSLVDKDL SFSEGLGSH FEFPDYCTPE 420  
 LSEMIAGDWL EANFSDLVFT Y

Seq ID NO: 658 DNA sequence  
 Nucleic Acid Accession #: NM\_001719  
 Coding sequence: 123..1418

10 1 11 21 31 41 51  
 | | | | | |  
 GGGCGCAGCG GGGCCCGTCT GCAGCAAGTG ACCGACGGCC GGGACGGCCG CCTGCCCCCT 60  
 CTGCCACCTG GGGCGGTGCG GGCCCGGAGC CCGGAGCCCG GGTAGCGCGT AGAGCCGGCG 120  
 CGATGCACGT GCGCTCACTG CGAGCTGCGG CGCCGCACAG CTTCGTGGCG CTCTGGGCAC 180  
 15 CCCTGTTCCCT GCTGCGCTCC GCCCTGGCCG ACTTCAGCCT GGACAACGAG GTGCACTCGA 240  
 GCTTCATCCA CCGGCGCCTC CGCAGCCAGG AGCGGCGGGA GATGCAGCGC GAGATCCTCT 300  
 CCATTTTGGG CTTGCCCCAC CGCCCGCGCC CGCACCTCCA GGGCAAGCAC AACTCGGCAC 360  
 CCATGTTTCT GCTGGACCTG TACAACGCCA TGGCGGTGGA GGAGGGCGGC GGGCCCGGCG 420  
 GCCAGGGCTT CTCCTACCCC TACAAGGCCG TCTTCAGTAC CCAGGGCCCC CCTCTGGCCA 480  
 20 GCCTGCAAGA TAGCCATTTC CTCACCGACG CCGACATGGT CATGAGCTTC GTCAACCTCG 540  
 TGGAACATGA CAAGGAATTC TTCCAGCTCA GCTACCACCA TCGAGAGTTC CGGTTTGATC 600  
 TTTCCAAGAT CCCAGAAGG GAAGCTGTCA CGGCGCCGA ATTCCGGATC TACAAGGACT 660  
 ACATCCGGGA ACGCTTCGAC AATGAGACGT TCCGGATCAG CGTTTATCAG GTGCTCCAGG 720  
 AGCACTTGGG CAGGGAATCG GATCTCTTCC TGCTCGACAG CCGTACCCTC TGGGCCTCGG 780  
 25 AGGAGGGCTG GCTGGTGTTC GACATCACAG CCACGAGCAA CCACTGGGTG GTCAATCCGC 840  
 GGCACAACCT GGGCCTGCAG CTCTCGGTGG AGACGCTGGA TGGGCAGAGC ATCAACCCCA 900  
 AGTTGGCGGG CCTGATTGGG CGGCACGGGC CCCAGAACAA GCAGCCCTTC ATGGTGGCTT 960  
 TCTTCAAGGC CACGGAGGTC CACTTCCGCA GCATCCGGTC CACGGGGAGC AAACAGCGCA 1020  
 GCCAGAACCG CTCCAAGACG CCAAGAACC AGGAAGCCCT GCGGATGGCC AACGTGGCAG 1080  
 30 AGAACAGCAG CAGCGACCAG AGGCAGGCCT GTAAGAAGCA CGAGCTGTAT GTCAGCTTCC 1140  
 GAGACCTGGG CTGGCAGGAC TGGATCATCG CGCCTGAAGG CTACGCCGCC TACTACTGTG 1200  
 AGGGGGAGTG TGCCTTCCCT CTGAACCTCT ACATGAACGC CACCAACCAC GCCATCGTGC 1260  
 AGACGCTGGT CCACTTCATC AACCCGGAAC CGGTGCCCAA GCCCTGCTGT GCGCCACGC 1320  
 AGCTCAATGC CATCTCCGTC CTCTACTTCG ATGACAGCTC CAACGTCATC CTGAAGAAAT 1380  
 35 ACAGAAACAT GGTGGTCCGG GCCTGTGGCT GCCACTAGCT CCTCCGAGAA TTCAGACCCT 1440  
 TTGGGGCCAA GTTTTCTGG ATCCTCCATT GCTCGCCTTG GCCAGGAACC AGCAGACCAA 1500  
 CTGCCTTTTG TGAGACCTTC CCCTCCCTAT CCCCAACTTT AAAGGTGTGA GAGTATTAGG 1560  
 AAACATGAGC AGCATATGGC TTTTGATCAG TTTTTCAGTG GCAGCATCCA ATGAACAAGA 1620  
 TCCTACAAGC TGTGCAGGCA AAACCTAGCA GGAAAAAACA ACAACGCATA AAGAAAAATG 1680  
 40 GCCGGGCCAG GTCAATTGGT GGAAGTGCAC AGCCATGCAC GGACTCGTTT CCAGAGGTAA 1740  
 TTATGAGCGC CTACAGGCA GGCCACCCAG CCGTGGGAGG AAGGGGGCGT GGCAAGGGGT 1800  
 GGGCACATTG GTGTCTGTGC GAAAGGAAAA TTGACCCGGA AGTTCCTGTA ATAAATGTCA 1860  
 CAATAAACG AATGAATG

Seq ID NO: 659 Protein sequence  
 Protein Accession #: NP\_001710

1 11 21 31 41 51  
 | | | | | |  
 50 MHVRSRLRAA PHSFVALWAP LFLRLSALAD FSLDNEVHSS FIHRLRSQRE RREMQREILS 60  
 ILGLPHRPRP HLQGHNSAP MFMLDLYNAM AVEEGGGPGG QGFSYPYKAV FSTQGPPLAS 120  
 LQDSHFLTDA DMVMSFVNLV EHDKEFFHPR YHHRERFDL SKIPEGEAVT AAEFRIYKDY 180  
 IRSDFLNETF RISVYQVLQE HLGRENDLFL LDSRTLWASE EGWLVDITA TSNHWVNNPR 240  
 55 HNLGLQLSVE TLDGQSINPK LAGLIGRHGP QNKQPFMVAE FKATEVHFRS IRSTGSKQRS 300  
 QNRSKTPKNQ EALRMANVAE NSSSDQRQAC KKHELYVSFR DLGWQDWIIA PEGYAAYYCE 360  
 GECAFLPLNSY MNATNHAIVQ TLVHFIPET VPKPCCAPTQ LNAISVLYFD DSSNVILKKY 420  
 RNMVVRACGC H

Seq ID NO: 660 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 211..1895

1 11 21 31 41 51  
 | | | | | |  
 65 GGATCTGAGG GGGCCCCAGT CACTTCCTCC ACGTTCTCGT GCTGGGCGGG AGGAGCGGAT 60  
 GGGGCTTGGG AGGCAGCCTG CTCTCCAGTC CCTATCCACC CACAGGTTTT TTGGGTCGGA 120  
 GAGGAATTAT CTGATAAAAT TCCTGGGTAA ATATTTTAA AAACGGAGAG TTTTAAAAA 180  
 TGATTTTTTT CCCTCGAAAA TGACCTTTTT ATGCTTCGAA GCAGTTTGTC AACCAGCATA 240  
 GTGCTTTTTT TTTTCTCTTC TTTTCTACG ATAAATGAAA GCATTTCTTC AAGAAAAAGG 300  
 70 CACAGGTTC TGAACAGCT GGATTCTGAT GGCACCATTA CTATAGAGGA GCAGATTGTC 360  
 CTTGTGCTGA AAGCGAAAGT ACAATGTGAA CTCACATCA CAGCTCAACT CCAGGAGGGA 420  
 GAAGGTAATT GTTCCCTGA ATGGGATGGA CTCATTGTT GGGCCAGAGG AACAGTGGGG 480  
 AAAATATCGG CTGTTCCATG CCCTCCTTAT ATTTATGACT TCAACCATAA AGGAGTTGCT 540  
 75 TTCCGACACT GTAACCCCAA TGGAACATGG GATTTTATGC ACAGCTTAAA TAAACATGG 600  
 GCCAATTATT CAGACTGCCT TCGCTTCTG CAGCCAGATA TCAGCATAGG AAAGCAAGAA 660  
 TTCTTTGAAC GCCTCTATGT AATGTATACC GTTGGCTACT CCATCTCTTT TGGTTCCTTG 720  
 GCTGTGGCTA TTCTCATCAT TGGTTACTTC AGACGATTGC ATTGCACTAG GAACTATATC 780  
 CACATGCAT TATTGTGTC TTTTCATCTG AGAGCTACAA GCATCTTTGT CAAAGACAGA 840  
 GTAGTCCATG CTCACATAGG AGTAAAGGAG CTGGAGTCCC TAATAATGCA GGATGACCCA 900  
 80 CAAAATTCCA TTGAGGCAAC TTCTGTGGAC AAATCACAAT ATATCGGGTG CAAGATTGCT 960  
 GTTGTGATGT TTATTACTT CCTGGCTACA AATTATTATT GGATCCTGGT GGAAGGTCTC 1020  
 TACCTGCATA ATCTCATCTT TGTGGCTTTC TTTTCGGACA CCAATACCT GTGGGGCTTC 1080  
 ATCTTGATAG GCTGGGGGT TCCAGCAGCA TTTGTTGCAG CATGGGCTGT GGCACGAGCA 1140  
 85 ACTCTGGCTG ATGCGAGGT CTGGAACCTT AGTCTGGAG ACATCAAGTG GATTATCAA 1200  
 GCACCGATCT TAGCAGCTAT TGGGCTGAAT TTTATTCTGT TTCTGAATAC GGTAGAGTT 1260  
 CTAGCTACCA AAATCTGGGA GACCAATGCA GTTGGGCATG ACACAAGGAA GCAATACAGG 1320  
 AAATGGCCA AATCGACACT GGTCTGTGTC CTAGTCTTTG GAGTGCATTA CATCGTGTTC 1380



GTATGCCTGC CTCATCCTT CACTGGGCTC GGGTGGGAGA TCCGCATGCA CTGTGAGCTC 1440  
 TTCTTCAACT CCTTTCAGGG TTTCTTTGTG TCTATCATCT ACTGCTACTG CAATGGAGAG 1500  
 GTTCAGGCAG AGGTGAAGAA GATGTGGAGT CCGTGGAAATC TCTCCGTGGA CTGGAAAAGG 1560  
 ACACCGCCAT GTGGCAGCCG CAGATGCGGC TCAGTGCTCA CCACCGTGAC GCACAGCACC 1620  
 AGCAGCCAGT CACAGGTGGC GGCCAGCACA CGCATGGTGC TTATCTCTGG CAAAGCTGCC 1680  
 AAGATCGCCA GCAGACAGCC TGACAGCCAC ATCACTTTAC CTGGCTATGT CTGGAGTAAC 1740  
 TCAGAGCAGG ACTGCCTGCC ACACCTCTTC CACGAGGAGA CCAAGGAAGA TAGTGGGAGG 1800  
 CAGGGAGATG ATATTCTAAT GGAGAAGCCT TCCAGGCCTA TGAATCTAA CCCAGACACT 1860  
 GAAGGATGCC AAGGAGAAAC TGAGGATGTT CTCTGA

Seq ID NO: 661 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | |  
 MLRSSLSSTSI VLFLESSFST INESSSRKR HRFLEQLDSD GTITIEEQIV LVLKAKVQCE 60  
 LNITAQLQEG EGNCPEWDG LICWPRGTVG KISAVPCPPY IYDFNHKGVA FRHCNPNGTW 120  
 DFMHSLNKTW ANYSDCLRFL QPDISIGKQE FFERLYVMYT VGYSISFGSL AVAILIIGYF 180  
 RRLHCTRNYY HMHLFVSFML RATSIFVKDR VVHAHIGVKE LESLIMQDDP QNSIEATSVD 240  
 KSQYIGCKIA VVMFIYFLAT NYYWILVEGL YLHNLIFVAF FSDTKYLWGF ILIGWGFPA 300  
 FVAAWAVARA TLADARCWEL SAGDIKWIYQ APILAAIGLN FILELNTVRV LATKIWETNA 360  
 VGHDTRKQYR KLAKSTLVLV LVFGVHYIVF VCLPHSFTGL GWEIRMHCEL FFNSFQGFV 420  
 SIIYCYCNGE VQAEVKMWS RWNLSVDWKR TTPCGSRRCG SVLTTVTHST SSQSQVAAS 480  
 RMVLISGKAA KIASRQPD SH ITPGYVWSN SEQDCLPHSF HEETKEDSGR QGDDILMEKP 540  
 SRPMESNPDT EGCQGETEDV L

Seq ID NO: 662 DNA sequence  
 Nucleic Acid Accession #: NM\_005048  
 Coding sequence: 143..1795

1 11 21 31 41 51  
 | | | | |  
 GGCCGGTGGC CCGGGCCCGA CCACCCAGC TGCGCGTCTG TACTGGCCAC AAGTTTGCTC 60  
 TGGGCCAGCC AAGTTGGCAA CTTGGAAGCT TCTCCCGGC TCTGGAGGAG GGTCCCTGCT 120  
 TCTTCCTACA GCCGTTCGG GCATGGCCGG GCTGGGGCG TCGCTCCACG TCTGGGGTTG 180  
 GCTAATGCTC GGCAGCTGCC TCCTGGCCAG AGCCAGCTG GATTCTGATG GCACCATTAC 240  
 TATAGAGGAG CAGATGTCC TTGTGCTGAA AGCGAAAGTA CAATGTGAAC TCAACATCAC 300  
 AGCTCAACTC CAGGAGGGAG AAGGTAATTG TTCCCTGAA TGGGATGGAC TCATTTGTTG 360  
 GCCCAGAGGA ACAGTGGGGA AAATATCGG TGTTCCATG CCTCCTTATA TTTATGACTT 420  
 CAACCATAAA GGAGTTGCTT TCCGACATG TAACCCCAAT GGAACATGGG ATTTTATGCA 480  
 CAGCTTAAAT AAAACATGGG CCAATTATTC AGACTGCCTT CGCTTTCTGC AGCCAGATAT 540  
 CAGCATAGGA AAGCAAGAAT TCTTTGAACG CCTCTATGTA ATGTATACCG TTGGCTACTC 600  
 CATCTCTTTT GGTTCCTTGG CTGTGGCTAT TCTCATCAT GGTACTTCA GACGATTGCA 660  
 TTGCACTAGG AACTATATCC ACATGCACTT ATTTGTGTCT TTCATGCTGA GAGCTACAAG 720  
 CATCTTTGTC AAAGACAGAG TAGTCCATGC TCACATAGGA GTAAAGGAGC TGGAGTCCCT 780  
 AATAATGCAG AATGACCCAC AAAATTCAT TGAGGCAACT TCTGTGGACA AATCACAATA 840  
 TATCGGGTGC AAGATTGCTG TTGTGATGTT TATTTACTTC CTGGCTACAA ATTATTATTG 900  
 GATCCTGGTG GAAGGTCTCT ACCTGCATAA TCTCATCTTT GTGGCTTTCT TTTCCGACAC 960  
 CAAATACCTG TGGGGCTTCA TCTTGATAGG CTGGGGGTTT CCAGCAGCAT TTGTTGCAGC 1020  
 ATGGGCTGTG GCACGAGCAA CTCTGGCTGA TGCGAGGTGC TGGGAACCTA GTGCTGGAGA 1080  
 CATCAAGTGG ATTTATCAAG CACCGATCTT AGCAGCTATT GGGCTGAATT TTATTCTGTT 1140  
 TCTGAATACG GTTAGAGTTC TAGCTACCAA AATCTGGGAG ACCAATGCAG TTGGGCATGA 1200  
 CACAAGGAAG CAATACAGGA AACTGGCCAA ATCGACACTG GTCTGTGCTC TAGTCTTTGG 1260  
 AGTGCATTAC ATCGTGTTCG TATGCCTGCC TCACTCCTTC ACTGGGCTCG GGTGGGAGAT 1320  
 CCGCATGCAC TGTGAGCTCT TCTTCAACTC CTTTCAGGGT TTCTTTGTGT CTATCATCTA 1380  
 CTGCTACTGC AATGGAGAGG TTCAGGCAGA GGTGAAGAAG ATGTGGAGTC GGTGGAATCT 1440  
 CTCCGTGGAC TGGAAAAGGA CACCGCCATG TGGCAGCCGC AGATGCGGCT CAGTGCTCAC 1500  
 CACCGTGACG CACAGCACCA GCAGCCAGTC ACAGGTGGCG GCCAGCACAC GCATGGTGCT 1560  
 TATCTCTGGC AAAGCTGCCA AGATCGCCAG CAGACAGCCT GACAGCCACA TCACTTTACC 1620  
 TGGCTATGTC TGGAGTAACT CAGAGCAGGA CTGCCTGCCA CACTCTTTCC ACGAGGAGAC 1680  
 CAAGGAAGAT AGTGGGAGGC AGGGAGATGA TATTCTAATG GAGAAGCCTT CCAGGCCTAT 1740  
 GGAATCTAAC CCAGACACTG AAGGATGCCA AGGAGAAACT GAGGATGTTT TCTGAATGGA 1800  
 CATTGTGGC TGACTTTTCA GGGCTGGTCC AATGGCTGGT TGTGTGAGAG GGCTTGGCTG 1860  
 ATACTCCTAT GCTTGAGTTC AAAGCTGAAA AATTGAGTTA AGGTGTTACT TAATAATAGT 1920  
 TTTTAGGCTC CATGAATTGG CTCCTGTAAA TACTAACGAC ATGAAAATGC AAGTGTCAAT 1980  
 GGAGTAGTTT ATTACCTTCT ATTGGCATCA AGTTTTCTCT TAAATTAATG TATGGTATTT 2040  
 GCTCTGTGAT TGTTCAATTT TTTCTGCTAC TTTTGGGTAG AAAAAAGATT CAATTGCTTG 2100  
 GCTGTAGCTT TCTCTCATAT ATATCACCTT AAATATAATG AAGATCTTTT AGTGTGTATC 2160  
 ATTTTCCTTT TAGAACTAG TATTCTCTTA TTTCTTACTT TAATGTACTT CTATCACTGC 2220  
 ATTTATTTTG CCTGTGCATA GGAGCAATTA GGATCTAAAA AAATATATGG GAAGATAAAA 2280  
 GATCTAAGAA CAAGTACTTG CTGGAAAATT AGTTGGCTGG ACATTGATAA AATAATGCAT 2340  
 TTATAACAAT TACATGTGTT TTTGGGAACA AGGAAAATTT CTCAAAAAAG AATATTTTAC 2400  
 ACATCCCTTC TTTTGAATGG CCTCTTTGTG ACCAGCCAGA CCTCAGGTCT TCACTCTTTC 2460  
 TTCTTTGTAA ACCATGTCAT GTGGAAAGAT TTCCTCAGTT AGTGAGCTTG TGTCTGCAAA 2520  
 TTGATTTTGT TTGTAATGTA TTTTGATAGC AAATCATGCT GCATCTATAT CTTTTTCTTG 2580  
 TTTGAGCTGT TACTACATTG TACATGGCAT GTGGGATCAA TTAAAAATTT GTTTTAAAAA 2640  
 T

Seq ID NO: 663 Protein sequence  
 Protein Accession #: NP\_005039

1 11 21 31 41 51  
 | | | | |  
 MAGLGASLHV WGWMLGSL LARAQLDSDG TITIEEQIVL VLKAKVQCEL NITAQLQEGE 60  
 GNCPEWDGL ICWPRGTVGK ISAVPCPPYI YDFNHKGVA FRHCNPNGTW FMSHSLNKTWA 120  
 NYSDCLRFLQ PDISIGKQEF FERLYVMYTV GYSISFGSLA VAILIIGYFR RLHCTRNYYH 180  
 MHLFVSFMLR ATSIFVKDRV VHAHIGVKEL ESLIMQDDPQ NSIEATSVDK SQYIGCKIAV 240



VMFIYFLATN YYWILVEGLY LHNLIFFVAFF SDTKYLWGF I LIGWGFPAAF VAAWAVARAT 300  
 LADARCWELS AGDIKWIYQA PILAAIGLNF ILFLNTVRVL ATKIWETNAV GHDTRKQYRK 360  
 LAKSTLVLVL VFGVHYIVFV CLPHSFTGLG WEIRMHCELF FNSFQGFVS IICYCNGEV 420  
 QAEVKKMWSR WNLSVDWKRT PPCGSRRCGS VLTTVTHSTS SQSQVAASR MVLISGKAAK 480  
 IASRQPDISHI TLPGYVWSNS EQDCLPHSFH EETKEDSGRQ GDDILMEKPS RPPMESNPDE 540  
 GCQGETEDVL

Seq ID NO: 664 DNA sequence  
 Nucleic Acid Accession #: NM\_012152  
 Coding sequence: 43..1104

1 11 21 31 41 51  
 | | | | |  
 CTTCTTTAAA TTTCTTTCTA GGATGTTTAC TTCTTCTCCA CAATGAATGA GTGTCACTAT 60  
 GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTCTG TGAAGTGGACA 120  
 GGAACAAAGC TTGTGATTGT TTTGTGTGTT GGGACGTTT TCTGCCTGTT TATTTTTTTT 180  
 TCTAATCTCT TGGTCATCGC GGCAGTGATC AAAACAGAA AATTTTCAAT CCCCTTCTAC 240  
 TACCTGTTGG CTAATTTAGC TGCTGCCGAT TTCTTCGCTG GAATTGCCTA TGTATTCCTG 300  
 ATGTTTAACA CAGGCCCACT TTCAAAACT TTGACTGTCA ACCGCTGGTT TCTCCGTCAG 360  
 GGGCTTCTGG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG 420  
 AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA 480  
 CTGCTCATTT TGCTTGTCTG GGCCATCGCC ATTTTTATGG GGGCGGTCCC CACACTGGGC 540  
 TGGAATTGCC TCTGCAACAT CTCTGCCTGC TCTTCCCTGG CCCCCATTTA CAGCAGGAGT 600  
 TACCTTGTTC TCTGGACAGT GTCCAACCTC ATGGCCTTCC TCATCATGGT TGTGGTGTAC 660  
 CTGCGGATCT ACGTGTACGT CAAGAGGAAA ACCAACGTCT TGTCTCCGCA TACAAGTGGG 720  
 TCCATCAGCC GCCGGAGGAC ACCCATGAAG CTAATGAAGA CCGTGATGAC TGTCTTAGGG 780  
 GCGTTTGTGG TATGCTGGAC CCCGGGCTGT GTGGTTCTGC TCCTCGACGG CCTGAACTGC 840  
 AGGCAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCCTGC TGCTGGCGCT GCTCAACTCC 900  
 GTCGTGAACC CCATCATCTA CTCCTACAAG GACGAGGACA TGTATGGCAG CATGAAGAAG 960  
 ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTCGCAT CCCCTCCACA 1020  
 GTCCTCAGCA GGAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080  
 GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCCTCTC GGCCCAACCA GGTGATGACT 1140  
 GTCTTAGG

Seq ID NO: 665 Protein sequence  
 Protein Accession #: NP\_036284

1 11 21 31 41 51  
 | | | | |  
 MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL VIAAVIKNRK 60  
 FHFPFYLLA NLAAADFFAG IAYVFLMNT GPVSKTLTVN RWFLRQGLLD SSLTASLTNL 120  
 LVIAVERHMS IMRMVHSNL TKKRVTLIL LVWAIIFMG AVPTLGWNCL CNISACSSLA 180  
 PIYSRSYLVF WTVSNLMAFL IMVVVYLRIV VYVKRKTNVL SPHTSGSISR RRTPMKLMKT 240  
 VMTVLGAFVV CWTPLVLVLL LDGLNCRQCG VQHVKRWFLL LALLNSVNP IIYSYKDEDM 300  
 YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS

Seq ID NO: 666 DNA sequence  
 Nucleic Acid Accession #: NM\_002821  
 Coding sequence: 150..3362

1 11 21 31 41 51  
 | | | | |  
 AACTCCCGCC TCGGGACGCC TCGGGGTCCG GCTCCGGCTG CGGCTGCTGC TCGGGCGCCC 60  
 GCGCTCCGGT GCGTCCGCTT CCTGTGCCCC CCGCGGAGCA GTCTGCGGCC CGCCGTGCGC 120  
 CCTCAGCTCC TTTTCTGAG CCCGCCGCGA TGGGAGCTGC GCGGGGATCC CCGGCCAGAC 180  
 CCCGCCGGTT GCCTCTGCTC AGCGTCCTGC TGCTGCCGCT GCTGGGCGGT ACCCAGACAG 240  
 CCATTGTCTT CATCAAGCAG CCGTCTCTCC AGGATGCACT GCAGGGGCGC CGGGCGCTGC 300  
 TTCGCTGTGA GGTGAGGCT CCGGGCCCGG TACATGTGTA CTGGCTGCTC GATGGGGCCC 360  
 CTGTCCAGGA CACGGAGCGG CGTTTCGCCC AGGGCAGCAG CCTGAGCTTT GCAGCTGTGG 420  
 ACCGGCTGCA GGAATCTGGC ACCTTCCAGT GTGTGGCTCG GGATGATGTC ACTGGAGAAG 480  
 AAGCCCGCAG TGCCAACGCC TCCTTCAACA TCAATGGAT TGAGGCAGGT CCTGTGGTCC 540  
 TGAAGCATCC AGCCTCGGAA GCTGAGATCC AGCCACAGAC CCAGGTCACT CTTCTGTTGCC 600  
 ACATTGATGG GCACCTCCGG CCCACCTACC AATGGTTCGG AGATGGGACC CCCCTTTCTG 660  
 ATGGTCAGAG CAACCACACA GTCAGCAGCA AGGAGCGGAA CCTGACGCTC CGGCCAGCTG 720  
 GTCCTGAGCA TAGTGGGCTG TATTCTGCTC GCGCCACAG TGCTTTTGGC CAGGCTTGCA 780  
 GCAGCCAGAA CTTACCTTG AGCATTGCTG ATGAAAGCTT TGCCAGGGTG GTGCTGGCAC 840  
 CCCAGGACGT GGTAGTAGCG AGGTATGAGG AGGCCATGTT CCATTGCCAG TTCTCAGCCC 900  
 AGCCACCCCC GAGCCTGCAG TGGCTCTTTG AGGATGAGAC TCCCATCACT AACCAGAGTC 960  
 GCGCCCGACA CCTCCGAGA GCCACAGTGT TTGCCAACGG GTCTCTGCTG CTGACCCAGG 1020  
 TCCGGCCACG CAATGCAGGG ATCTACCGCT GCATTGGCCA GGGGCAGAGG GGCCCAACCA 1080  
 TCATCTTGA AGCCACACTT CACCTAGCAG AGATTGAAGA CATGCCGCTA TTTGAGCCAC 1140  
 GGGTGTTTAC AGCTGGCAGC GAGGAGCGTG TGACCTGCC TCCCCCAAG GGTCTGCCAG 1200  
 AGCCAGCGT GTGGTGGGAG CACGCGGGAG TCCGGCTGCC CACCCATGGC AGGGTCTACC 1260  
 AGAAGGGCCA CGAGCTGGTG TTGGCCAATA TTGCTGAAAG TGATGCTGGT GTCTACACCT 1320  
 GCCACGCGGC CAACCTGGCT GGTGAGCGGA GACAGGATGT CAACATCACT GTGGCCACTG 1380  
 TGCCCTCCTG GCTGAAGAAG CCCCAAGACA GCCAGCTGGA GGAGGGCAA CCCGGCTACT 1440  
 TGGATTGCCT GACCCAGGCC ACACCAAAAC CTACGATTGT CTGGTACAGA AACCAGATGC 1500  
 TCATCTAGA GGACTACGG TTCGAGGTCT TCAAGAATGG GACCTTGCGC ATCAACAGCG 1560  
 TGGAGGTGTA TGATGGGACA TGGTACCGTT GTATGAGCAG CACCCAGGCC GGCAGCATCG 1620  
 AGGCGCAAGC CCGTGTCCAA GTGCTGAAA AGCTCAAGTT CACACCACCA CCCCAGCCAC 1680  
 AGCAGTGCAT GGAGTTTGAC AAGGAGGCCA CCGTGCCCTG TTCAGCCACA GGCCGAGAGA 1740  
 AGCCCACTAT TAAGTGGGAA CGGGCAGATG GGAGCAGCCT CCCAGAGTGG GTGACAGACA 1800  
 ACGCTGGGAC CTGCAATTTT GCGCGGGTGA CTCGAGATGA CGCTGGCAAC TACACTTGCA 1860  
 TTGCCTCCAA CCGGCCGAG GTGCCAGATT GTGCCATGT CCAGCTCACT GTGGCAGTTT 1920  
 TTATCACCTT CAAAGTGGAA CCAGAGCGTA CGACTGTGTA CCAGGGCCAC ACAGCCCTAC 1980  
 TGCAGTGCAG GGGCCAGGGG GACCCCAAGC CGCTGATTCA GTGGAAAGGC AAGGACCGCA 2040  
 TCCTGGACCC CACCAAGCTG GGACCCAGGA TGCACATCTT CCAGAAATGG TCCCTGGTGA 2100



	TCCATGACGT	GGCCCCCTGAG	GACTCAGGCC	GCTACACCTG	CATTGCAGGC	AACAGCTGCA	2160
	ACATCAAGCA	CACGGAGGCC	CCCCTCTATG	TCGTGGACAA	GCCTGTGCCG	GAGGAGTCGG	2220
	AGGGCCCTGG	CAGCCCTCCC	CCCTACAAGA	TGATCCAGAC	CATTGGGTTG	TCGGTGGGTG	2280
5	CCGCTGTGGC	CTACATCATT	GCCGTGCTGG	GCCTCATGTT	CTACTGCAAG	AAGCGCTGCA	2340
	AAGCCAAGCG	GCTGCAGAAG	CAGCCCCGAG	GCGAGGAGCC	AGAGATGGAA	TGCCTCAACG	2400
	GAGGGCCTTT	GCAGAACGGG	CAGCCCTCAG	CAGAGATCCA	AGAAGAAGTG	GCCTTGACCA	2460
	GCTTGGGCTC	CGGCCCCGCG	GCCACCAACA	AACGCCACAG	CACAAGTGAT	AAGATGCACT	2520
	TCCCACGGTC	TAGCCTGCAG	CCCATCACCA	CGCTGGGGAA	GAGTGAGTTT	GGGGAGGTGT	2580
10	TCCTGGCAAA	GGCTCAGGGC	TTGGAGGAGG	GAGTGGCAGA	GACCCCTGGT	CTTGTGAAGA	2640
	GCCTGCAGAC	GAAGGATGAG	CAGCAGCAGC	TGGACTTCCG	GAGGGAGTTG	GAGATGTTTG	2700
	GGAAGCTGAA	CCACGCCAAC	GTGGTGCGGC	TCCTGGGGCT	GTGCCGGGAG	GCTGAGCCCC	2760
	ACTACATGGT	GCTGGAATAT	GTGGATCTGG	GAGACCTCAA	GCAGTTCCTG	AGGATTTCCA	2820
	AGAGCAAGGA	TGAAAAATTG	AAGTCACAGC	CCCTCAGCAC	CAAGCAGAAG	GTGGCCCTAT	2880
15	GCACCCAGGT	AGCCCTGGGC	ATGGAGCACC	TGTCCAACAA	CCGCTTTGTG	CATAAGGACT	2940
	TGGCTGCGCG	TAACTGCCTG	GTCAGTGCCC	AGAGACAAAG	GAAGGTGTCT	GCCCTGGGCC	3000
	TCAGCAAGGA	TGTGTACAAC	AGTGAGTACT	ACCACTTCCG	CCAGGCCTGG	GTGCCGCTGC	3060
	GCTGGATGTC	CCCCGAGGCC	ATCCTGGAGG	GTGACTTCTC	TACCAAGTCT	GATGTCTGGG	3120
	CCTTCGGTGT	GCTGATGTGG	GAAGTGTTTA	CACATGGAGA	GATGCCCCAT	GGTGGGCAGG	3180
20	CAGATGATGA	AGTACTGGCA	GATTTGCAGG	CTGGGAAGGC	TAGACTTCCT	CAGCCCGAGG	3240
	GCTGCCCTTC	CAAACTCTAT	CGGCTGATGC	AGCGCTGCTG	GGCCCTCAGC	CCCAAGGACC	3300
	GGCCCTCCTT	CAGTGAGATT	GCCAGCGCCC	TGGGAGACAG	CACCGTGGAC	AGCAAGCCGT	3360
	GAGGAGGGAG	CCCGCTCAGG	ATGGCCTGGG	CAGGGGAGGA	CATCTCTAGA	GGGAAGCTCA	3420
	CAGCATGATG	GGCAAGATCC	CTGTCTCTCT	GGGCCCTGAG	GTGCCCTAGT	GCAACAGGCA	3480
25	TTGCTGAGGT	CTGAGCAGGG	CCTGGCCTTT	CCTCCTCTTC	CTCACCTCA	TCCTTTGGGA	3540
	GGCTGACTTG	GACCCAACT	GGGCGACTAG	GGCTTTGAGC	TGGGCAGTTT	CCCCTGCCAC	3600
	CTCTTCTCT	ATCAGGGACA	GTGTGGGTGC	CACAGGTAAC	CCCAATTTCT	GGCCTTCAAC	3660
	TTCTCCCTT	GACCGGGTCC	AACTCTGCCA	CTCATCTGCC	AACTTTGCCT	GGGGAGGGCT	3720
	AGGCTTGGGA	TGAGCTGGGT	TTGTGGGGAG	TTCTTAATA	TTCTCAAGTT	CTGGGCACAC	3780
30	AGGGTTAATG	AGTCTCTTGC	CCACTGGTCC	ACTTGGGGGT	CTAGACCAGG	ATTATAGAGG	3840
	ACACAGCAAG	TGAGTCTCTC	CCACTCTGGG	CTTGTGCACA	CTGACCCAGA	CCCACGTCTT	3900
	CCCCACCTT	CTCTCCTTTC	CTCATCTTAA	GTGCCTGGCA	GATGAAGGAG	TTTTCAGGAG	3960
	CTTTTGACAC	TATATAAACC	GCCCTTTTGG	TATGCACCAC	GGGCGGCTTT	TATATGTAAT	4020
	TGCAGCGTGG	GGTGGGTGGG	CATGGGAGGT	AGGGGTGGGC	CCTGGAGATG	AGGAGGGTGG	4080
35	GCCATCCTTA	CCCCACACTT	TTATTGTTGT	CGTTTTTTGT	TTGTTTTGTT	TTTTTGTTTT	4140
	TGTTTTTGT	TTTACACTCG	CTGCTCTCAA	TAAATAAGCC	TTTTTTTA		

Seq ID NO: 667 Protein sequence  
Protein Accession #: NP\_002812

40	1	11	21	31	41	51	
	MGAARGSPAR	PRRLPLLSVL	LLPLLGGTQT	AIVFIKQPSS	QDALQGRRAL	LRCEVEAPGP	60
	VHVVWLLDGA	PVQDTERRFA	QGSSLSFAAV	DRLQDSGTFQ	CVARDDVTGE	EARSANASFN	120
45	IKWIEAGPVV	LKHPASEAEI	QPQTQVTLRC	HIDGHPRTY	QWFRDGTPLS	DGQSNHTVSS	180
	KERNLTLRPA	GPEHSGLYSC	CAHSAGQAC	SSQNFSLIA	DESFAFVLA	PQDVVARYE	240
	EAMFHCQFSA	QPPPSLQWLF	EDETPTNRS	RPPHLRRATV	FANGSLLLTQ	VRPRNAGIYR	300
	CIGQGQRGPP	IILEATLHLA	EIEDMPLFEP	RVFTAGSEER	VTCLPPKGLP	EPSVWWEHAG	360
	VRLPTHGRVY	QKGHELVLAN	IAESDAGVYT	CHAANLAGQR	RQDVNITVAT	VPSWLKKPQD	420
50	SQLEEGKPGY	LDCLTQATPK	PTVVWYRNQM	LISEDSTRFEV	FKNGTLRINS	VEVYDGTWYR	480
	CMSSTPAGSI	EAQARVQVLE	KLKFTPPPQP	QQCMFEDKEA	TVPCSATGRE	KPTIKWERAD	540
	GSSLPEWVTD	NAGTLHFARV	TRDDAGNYTG	IASNGPQGI	RAHVQLTVAV	FITFKVEPER	600
	TTVYQGHTAL	LQCEAQGDPK	PLIQWKGKDR	ILDPTKLGP	MHIFQNGSLV	IHDVAPEDSG	660
	RYTCIAGNSC	NIKHTEAPLY	VVDKPVPEES	EGPGSPPPYK	MIQTIGLSVG	AAVAYIIAVL	720
55	GLMFYCKKRC	KAKRLQKQPE	GEPEMECLN	GGPLQNGQPS	AEIQEEVALT	SLGSGPAATN	780
	KRHSTSDKMH	FPRSSLQPIT	TLGKSEFGEV	FLAKAQGLEE	GVAETLVLVK	SLQTKDEQQQ	840
	LDFRRELEMF	GKLNHANVVR	LLGLCREAEP	HYMVLEYVDL	GDLKQFLRIS	KSKDEKLKSQ	900
	PLSTKQKVAL	CTQVALGMEH	LSNNRFVHKD	LAARNCLVSA	QRQVKVSAAG	LSKDVYNSEY	960
	YHFRQAWVPL	RWMSPEAILE	GDFSTKSDVW	AFGLVMWEVF	THGEMPHGGQ	ADDEVLDLQ	1020
60	AGKARLPQPE	GCPSKLYRLM	QRCWALSPKD	RPSFSEIASA	LGDSTVDSKP		

Seq ID NO: 668 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1389

65	1	11	21	31	41	51	
	ATGGGCTACC	AGAGGCAGGA	GCCTGTCATC	CCGCCGCAGA	GAGATTTAGA	TGACAGAGAA	60
	ACCTTGTGTT	CTGAACATGA	GTATAAAGAG	AAAACCTGTC	AGTCTGCTGC	TCTTTTAAAT	120
70	GTTGTCAACT	CGATTATAGG	ATCTGGTATA	ATAGGATTGC	CTTATTCAAT	GAAGCAAGCT	180
	GGGTTTCCTT	TGGGAATATT	GCTTTTATTC	TGGGTTTCAT	ATGTTACGGA	CTTTTCCCTT	240
	GTTTTATTGA	TAAAAGGAGG	GGCCCTCTCT	GGAACAGATA	CCTACCAGTC	TTTGGTCAAT	300
	AAAACTTTCG	GCTTTCAGG	GTATCTGCTC	CTCTCTGTTC	TTCAGTTTTT	GTATCCTTTT	360
	ATAGCAATGA	TAAATTACAA	TATAATAGCT	GGAGATACTT	TGAGCAAAGT	TTTTCAAAGA	420
75	ATCCCAGGAG	TTGATCCTGA	AAACGTGTTT	ATTGGTCGCC	ACTTCATTAT	TGGACTTTCC	480
	ACAGTTACCT	TTACTCTGCC	TTTATCCTTG	TACCGAATA	TAGCAAAGCT	TGGAAAGGTC	540
	TCCCTCACTT	CTACAGTTT	AACACTCTG	ATCTCTGGAA	TTGTAATGGC	AAGGGCAATT	600
	TCACTGGGTC	CACACATACC	AAAAACAGAA	GACGCTTGGG	TATTTGCAAA	GCCCAATGCC	660
	ATTCAAGCGG	TCGGGGTTAT	GTCTTTTGCA	TTTATTGGCC	ACCATAACTC	CTTCTTAGTT	720
80	TACAGTTCTC	TAGAAGAACC	CACAGTAGCT	AAGTGGTCCC	GCCTTATCCA	TATGTCCATC	780
	GTGATTTCTG	TATTTATCTG	TATATCTCTT	GCTACATGTG	GATACTTGAC	ATTTACTGGC	840
	TTCAACCAAG	GGGACTTATT	TGAAAATTAC	TGCAGAAATG	ATGACCTGGT	AACATTTGGA	900
	AGATTTTGT	ATGGTGTAC	TGTCAATTTG	ACATACCTTA	TGGAATGCTT	TGTGACAAGA	960
	GAGGTAATTG	CCAATGTGTT	TTTTGGTGGG	AATCTTTCAT	CGGTTTTCCA	CATTGTTGTA	1020
85	ACAGTGATGG	TCATCACTGT	AGCCACGCTT	GTGTCAATGC	TGATTGATTG	CCTCGGGATA	1080
	GTTCTAGAAC	TCAATGGTGT	GCTCTGTGCA	ACTCCCTCA	TTTTTATCAT	TCCATCAGCC	1140
	TGTTATCTGA	AACTGTCTGA	AGAACCAAGG	ACACACTCCG	ATAAGATTAT	GTCTTGTGTC	1200
	ATGCTTCCCA	TTGGTGCTGT	GGTGATGGTT	TTTGGATTCC	TCATGGCTAT	TACAAATACT	1260



CAAGACTGCA CCCATGGGCA GGAAATGTTT TACTGCTTTC CTGACAATTT CTCTCTCACA 1320  
 AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACTTT CTACTTTAA TATTAGTATC 1380  
 TTTCAATGA

5 Seq ID NO: 669 Protein sequence  
 Protein Accession #: Eos sequence

10 1 11 21 31 41 51  
 | | | | | |  
 MGYQRQEPVI PPQRDLDDRE TLVSEHEYKE KTCQSAALFN VVNSIIGSGI IGLPYSMKQA 60  
 GFPLGILLLF WVSIVTDFSL VLLIKGGALS GTDLYQSLVN KTFGFPYLL LSVLQFLYPF 120  
 IAMISYNIIA GDTLSKVFR IPGVDPENVF IGRHFIIGLS TVTFTLPLSL YRNIKLGKV 180  
 SLISTGLTTL ILGIVMARAI SLGPHIPKTE DAWVFAKPNA IQAVGVMSFA FICHNSFLV 240  
 YSSLEETVA KWSRLIHMSI VISVFICIFF ATCGYLTFTG FTQGLDFENY CRNDDLVTFG 300  
 15 RFCYGVTVIL TYPMECFVTR EVIANVFFGG NLSSVFHIVV TVMVITVATL VSLLDCLGI 360  
 VLELNGVLCA TPLIFIIPSA CYLKLSEEP THSDKIMSCV MLPIGAVMV FGFVMAITNT 420  
 QDCTHQEMF YCFPDNFSLT NTSESHVQQT TQLSTLNISI FQ

20 Seq ID NO: 670 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..1284

25 1 11 21 31 41 51  
 | | | | | |  
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGGATTGCC TTATTCAATG 60  
 AAGCAAGCTG GGTTCCTTT GGAATATTG CTTTATCTCT GGGTTTCATA TGTTACAGAC 120  
 TTTTCCCTTG TTTTATTGAT AAAAGGAGGG GCCCTCTCTG GAACAGATAC CTACCAGTCT 180  
 TTGGTCAATA AAACCTTCGG CTTTCCAGGG TATCTGCTCC TCTCTGTTCT TCAGTTTTTG 240  
 TATCCTTTTA TAGCAATGAT AAGTTACAAT ATAATAGCTG GAGATACTTT GAGCAAAGTT 300  
 30 TTTCAAAGAA TCCCAGGAGT TGATCCTGAA AACGTGTTTA TTGGTCGCCA CTTTATTATT 360  
 GGACTTTCCA CAGTTACCTT TACTCTGCCT TTATCCTTGT ACCGAAATAT AGCAAAGCTT 420  
 GGAAAGGTCT CCTCATCTC TACAGGTTTA ACAACTCTGA TTCTTGGAAT TGTAATGGCA 480  
 AGGGCAATTT CACTGGGTCC ACACATACCA AAAACAGAAG ACGCTTGGGT ATTTGCAAAG 540  
 CCCAATGCCA TTCAAGCGGT CGGGGTTATG TCTTTTGCAT TTATTTGCCA CCATAACTCC 600  
 35 TTCTTAGTTT ACAGTTCTCT AGAAGAACCC ACAGTAGCTA AGTGGTCCCG CCTTATCCAT 660  
 ATGTCCATCG TGATTTCTGT ATTTATCTGT ATATCTTTG CTACATGTGG ATACTTGACA 720  
 TTTACTGGCT TCACCAAGG GGACTTATTT GAAATTAAT GCAGAAATGA TGACCTGGTA 780  
 ACATTTGGA GATTTTGTTA TGGTGTCACT GTCATTTGA CATACCCTAT GGAATGCTTT 840  
 GTGACAAGAG AGGTAATTGC CAATGTGTTT TTTGTGGGA ATCTTTCATC GGTTTTCCAC 900  
 40 ATTGTTGTAA CAGTGATGGT CATCACTGTA GCCACGCTTG TGTCAATGCT GATTGATTGC 960  
 CTCGGGATAG TTCTAGAAT CAATGGTGTG CTCTGTGCAA CTCCCTCAT TTTTATCATT 1020  
 CCATCAGCCT GTTATCTGAA ACTGTCTGAA GAACCAAGGA CACACTCCGA TAAGATTATG 1080  
 TCTTGTGCA TGCTTCCCAT TGGTGTCTGT GTGATGGTTT TTGGATTCTG CATGGCTATT 1140  
 45 ACAAATACTC AAGACTGCAC CCATGGGCAG GAAATGTCT ACTGCTTTCC TGACAATTTT 1200  
 TCTCTCACA ATACCTCAGA GTCTCATGTT CAGCAGACAA CACAACCTTC TACTTTAAAT 1260  
 ATTAGTATCT TTCAACTCGA GTAA

50 Seq ID NO: 671 Protein sequence  
 Protein Accession #: Eos sequence

55 1 11 21 31 41 51  
 | | | | | |  
 MGYQRQEPVI PPQRGLPYSM KQAGFPLGIL LLEWVSIVTD FSLVLLIKGG ALSGTDYQS 60  
 LVNKTGFPFG YLLLSVLQFL YPFIAMISYN IIAGDTLSKV FQRIQGVDP NVFIGRHFII 120  
 GLSTVTFTLP LSLYRNIKAL GKVSISTGL TLLILGIVMA RAISLGHIP KTEDAWVFAK 180  
 PNAIQAVGVM SFAFICHNS FLVYSSLEEP TVAKWSRLIH MSIVISVFIC IFFATCGYLT 240  
 FTGFTQGLDF ENYCRNDDL TFGRFYCVGT VILTYPMECF VTREVIANVF FGNLSSVFH 300  
 IVVTVMVITV ATLVSLLIDC LGIVLELNGV LCATPLIFII PSACYLKLSE EPRTHSDKIM 360  
 60 SCVMLPIGAV VMVFGFVMAI TNTQDCTHQG EMFYCFPDNF SLTNTSESHV QOTTQLSTLN 420  
 ISIFQLE

65 Seq ID NO: 672 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..1203

70 1 11 21 31 41 51  
 | | | | | |  
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGT TTTCCCTTGT TTTATTGATA 60  
 AAAGGAGGGG CCTCTCTGG AACAGATACC TACAGTCTT TGGTCAATAA AACTTTCGGC 120  
 TTTCCAGGGT ATCTGCTCCT CTCTGTTCTT CAGTTTTTGT ATCCTTTTAT AGCAATGATA 180  
 AGTTACAATA TAATAGCTGG AGATACTTTG AGCAAAGTTT TTCAAAGAAT CCCAGGAGTT 240  
 GATCCTGAAA ACGTGTATTAT TGGTCGCCAC TTCAATTATG GACTTTCCAC AGTTACCTTT 300  
 ACTCTGCCTT TATCCTTGTA CCGAAATATA GCAAAGCTTG GAAAGGTCTC CCTCATCTCT 360  
 75 ACAGGTTTAA CAACTCTGAT TCTTGGAATT GTAATGGCAA GGGCAATTTT ACTGGGTCCA 420  
 CACATACCAA AAACAGAAGA CGCTTGGGTA TTTGCAAAGC CCAATGCCAT TCAAGCGGTC 480  
 GGGGTTATGT CTTTGGCATT TATTTGCCAC CATAACTCCT TCTTAGTTTA CAGTTCTCTA 540  
 GAAGAACCCA CAGTAGCTAA GTGGTCCCGC CTTATCCATA TGTCCATCGT GATTTCTGTA 600  
 TTTATCTGTA TATTCTTTGC TACATGTGGA TACTTGACAT TTAGTGGCTT CACCCAAGGG 660  
 GACTTATTTG AAAATTACTG CAGAAATGAT GACCTGGTAA CATTTGGAAG ATTTTGTATT 720  
 80 GGTGTCACTG TCATTTTGAC ATACCCTATG GAATGCTTTG TGACAAGAGA GGTAATTGCC 780  
 AATGTGTTTT TTGGTGGGAA TCTTTCATCG GTTTTCCACA TTGTTGTAAC AGTGATGGTC 840  
 ATCAGTGTAG CCAGCTTGT GTCAATTGCT ATTGATTGCC TCGGGATAGT TCTAGAAGTC 900  
 AATGGTGTGC TCTGTGCAAC TCCCCTCATT TTTATCATTC CATCAGCCTG TTATCTGAAA 960  
 CTGTCTGAAG AACCAAGGAC AACTCCGAT AAGATTATGT CTTGTGTCAT GCTTCCCAT 1020  
 85 GGTGCTGTGG TGATGGTTTT TGGATTCTGC ATGGCTATTA CAAATACTCA AGACTGCACC 1080  
 CATGGGCAGG AATGTTCTA CTGCTTTCCT GACAATTTCT CTCTCAGAAA TACCTCAGAG 1140  
 TCTCATGTTC AGCAGACAAC ACAACTTCTT ACTTTAAATA TTAGTATCTT TCAACTCGAG 1200



TAA

Seq ID NO: 673 Protein sequence  
Protein Accession #: Eos sequence

```

1      11      21      31      41      51
|      |      |      |      |      |
MGYQRQEPVI PPQFSLVLLI KGGALSGTDT YQSLVNKTFG FPGYLLLSVL QFLYPFIAMI 60
SYNIIAGDTL SKVFQRIPIV DPENVFGRH FIIGLSTVTF TLPLSLYRNI AKLGKVSLLS 120
TGLTTLILGI VMARAIISLP HIPKTEDAWV FAKPNAIQAV GVMSFAFICH HNSFLVYSSL 180
EPTVAKWSR LIHMSIVISV FICIFFATCG YLFTGTQDQ DLFENYCRND DLVTGFRFCY 240
GVTVILTYPM ECFVTREVI NVFFGGNLSS VFHIVVTVMV ITVATLVSLI IDCLGIVLEL 300
NGVLCATPLI FIIPSACYLK LSEEPRTSD KIMSCVMLPI GAVVMVFGFV MAITNTQDCT 360
HGQEMFYCFP DNFSLTNTSE SHVQQTQLS TLNISIFQLE

```

Seq ID NO: 674 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1140

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGAGG TCAATAAAAC TTTCGGCTTT 60
CCAGGGTATC TGCTCCTCTC TGTTCTTCAG TTTTGTATC CTTTATAGC AATGATAAGT 120
TACAATATAA TAGCTGGAGA TACTTTGAGC AAAGTTTTTC AAAGAATCCC AGGAGTTGAT 180
CCTGAAAACG TGTATTATTG TCGCCACTTC ATTATTGGAC TTTCACAGT TACCTTTACT 240
CTGCCTTTAT CCTGTACCG AAATATAGCA AAGCTTGGAA AGGTCTCCCT CATCTCTACA 300
GGTTTAACAA CTCTGATTCT TGGAAATGTA ATGGCAAGGG CAATTTCACT GGGTCCACAC 360
ATACCAAAAA CAGAAGACGC TTGGGTATTT GCAAAGCCCA ATGCCATTCA AGCGGTCGGG 420
GTTATGTCTT TTGCATTTAT TTGCCACCAT AACTCCTTCT TAGTTTACAG TTCTCTAGAA 480
GAACCCACAG TAGCTAAGTG GTCCCGCCTT ATCCATATGT CCATCGTGAT TTCTGTATTT 540
ATCTGTATAT TCTTTGCTAC ATGTGGATAC TTGACATTTA CTGGCTTCAC CCAAGGGGAC 600
TTATTTGAAA ATTACTGCAG AAATGATGAC CTGGTAACAT TTGGAAGATT TTGTTATGGT 660
GTCACCTGTA TTTTGACATA CCCTATGGAA TGCTTTGTGA CAAGAGAGGT AATTGCCAAT 720
GTGTTTTTTG GTGGGAATCT TTCATCGGTT TTCCACATTG TTGTAACAGT GATGGTCATC 780
ACTGTAGCCA CGCTTGTGTC ATTGCTGATT GATTGCCTCG GGATAGTTCT AGAACTCAAT 840
GGTGTGCTCT GTGCAACTCC CCTCATTTTT ATCATTCAT CAGCCTGTTA TCTGAAACTG 900
TCTGAAGAAC CAAGGACACA CTCCGATAAG ATTATGTCTT GTGTCATGCT TCCCATTTGGT 960
GCTGTGGTGA TGGTTTTTGG ATTCGTCATG GCTATTACAA ATACTCAAGA CTGCACCCAT 1020
GGGCAGGAAA TGTTCTACTG CTTTCCTGAC AATTCTCTC TCACAAATAC CTCAGAGTCT 1080
CATGTTTCAGC AGACAACACA ACTTTCTACT TTAAATATTA GTATCTTTCA ACTCGAGTAA

```

Seq ID NO: 675 Protein sequence  
Protein Accession #: Eos sequence

```

1      11      21      31      41      51
|      |      |      |      |      |
MGYQRQEPVI PPQVNKTFGF PGYLLLSVLQ FLYPFIAMIS YNIIAGDTLS KVFQRIPIVD 60
PENVFGRHF IIGLSTVTF LPLSLYRNIA KLKVSLLIST GLTTLILGIV MARAIISLPH 120
IPKTEDAWVF AKPNAIQAVG VMSFAFICH NSFLVYSSLE EPTVAKWSRL IHMSIVISVF 180
ICIFFATCGY LTFTGTQGD LFNENYCRND LVTFGRFCY VTVILTYPME CFVTREVIAN 240
VFVGGNLSSV FHIVVTVMVI TVATLVLELN GVLCAATPLI IIPSACYLKL 300
SEEPRTSDK IMSCVMLPIG AVVMVFGFV AITNTQDCTH GQEMFYCFPD NFSLTNTSES 360
HVQQTQLST LNISIFQLE

```

Seq ID NO: 676 DNA sequence  
Nucleic Acid Accession #: NM\_006853.1  
Coding sequence: 26..874

```

1      11      21      31      41      51
|      |      |      |      |      |
AGGAATCTGC GCTCGGGTTC CGCAGATGCA GAGGTTGAGG TGGCTGCGGG ACTGGAAGTC 60
ATCGGGCAGA GGTCTCACAG CAGCCAAGGA ACCTGGGGCC CGCTCCTCCC CCCTCCAGGC 120
CATGAGGATT CTGCAGTTAA TCCTGCTTGC TCTGGCAACA GGGCTTGTAG GGGGAGAGAC 180
CAGGATCATC AAGGGGTTCC AGTGCAAGCC TCACTCCCAG CCCTGGCAGG CAGCCCTGTT 240
CGAGAAGACG CGGCTACTCT GTGGGGCGAC GCTCATCGCC CCCAGATGGC TCCTGACAGC 300
AGCCCACTGC CTCAAGCCCC GCTACATAGT TCACCTGGGG CAGCACAACC TCCAGAAGGA 360
GGAGGGCTGT GAGCAGACCC GGACAGCCAC TGAGTCCTTC CCCCACCCCG GCTTCAACAA 420
CAGCCTCCCC AACAAAGACC ACCGCAATGA CATCATGCTG GTGAAGATGG CATCGCCAGT 480
CTCCATCACC TGGGCTGTGC GACCCCTCAC CCTCTCCTCA CGCTGTGTCA CTGCTGGCAC 540
CAGCTGCCTC ATTTCCGGCT GGGGCAGCAC DCLGIVLELN CAGTTACGCC TGCCTCACAC 600
CTTGCGATGC GCCAATATCA CCATCATTTG GCACCAAGG TGTGAGAAGC CCTACCCCGG 660
CAACATCACA GACACCATGG TGTGTGCCAG CGTGCAGGAA GGGGGCAAGG ACTCCTGCCA 720
GGGTGACTCC GGGGGCCCTC TGGTCTGTAA CCAGTCTCTT CAAGGCATTA TCTCCTGGGG 780
CCAGGATCCG TGTGCGATCA CCCGAAAGCC TGGTGTCTAC ACGAAAGTCT GCAAATATGT 840
GGACTGGATC CAGGAGACGA TGAAGAACAA TTAGACTGGA CCCACCCACC ACAGCCCATC 900
ACCTCCCAT TCCACTTGGT GTTTGGTTCC TGTACTCTCT GTTAATAAGA AACCTAAGC 960
CAAGACCCCT TACGAACATT CTTTGGGCCT CTTGGACTAC AGGAGATGCT GTCACCTAAT 1020
AATCAACCTG GGGTTCGAAA TCAGTGAGAC CTGGATTCAA ATTCTGCCTT GAAATATTGT 1080
GACTCTGGGA ATGACAACAC CTGGTTTGTG CTCTGTGTA TCCCCAGCCC CAAAGACAGC 1140
TCCTGGCCAT ATATCAAGGT TTCAATAAAT ATTTGCTAAA TGAGTG

```

Seq ID NO: 677 Protein sequence  
Protein Accession #: NP\_006844.1

```

1      11      21      31      41      51
|      |      |      |      |      |
MRILQLILLA LATGLVGGET RIIKGFECKP HSQPWQAALF EKTRLLCGAT LIAPRWLLTA 60

```



AHCLKPRYIV HLGQHNLOKE EGCEQTRTAT ESFPHPGFNN SLPNKDHRND IMLVKMASPV 120  
SITWAVRPLT LSSRCVTAGT SCLISGWGST SSPQLRLPHT LRCANITIE HQKCENAYPG 180  
NITDTMVCAS VQEGGKDSQ GDSGGPLVCN QSLQGIISWG QDPCAITRKP GVYTKVCKYV 240  
DWIQETMKNN

Seq ID NO: 678 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..933

1 11 21 31 41 51  
| | | | | |  
ATGTGCAGCA ATGGACGGTG CATCCCGGGC GCCTGGCAGT GTGACGGGCT GCCTGACTGC 60  
TTCGACAAGA GTGATGAGAA GGAGTGCCCC AAGGCTAAGT CGAAATGTGG CCCGACCTTC 120  
TCCCCCTGTG CCAGCGGCAT CCATTGCATC ATTGGTCGCT TCCGGTGCAA TGGGTTTGAG 180  
GACTGTCCCG ATGGCAGCGA TGAAGAGAAC TGCACAGCAA ACCCTCTGCT TGCTCCACC 240  
GCCCCTACC ACTGCAAGAA CGGCCTCTGT ATTGACAAGA GCTTCATCTG CGATGGACAG 300  
AATAACTGTC AAGACAACAG TGATGAGGAA AGCTGTGAAA GTTCTCAAGA ACCCGGCAGT 360  
GGGACAGGTG TTGTGACTTC AGAGAACCAA CTTGTGTATT ACCCCAGCAT CACCTATGCC 420  
ATCATCGGCA GCTCCGTCAT TTTTGTGCTG GTGGTGCCCT TGCTGGCACT GGTCTTGCAC 480  
CACCAGCGGA AGCGGAACAA CCTCATGACG CTGCCCGTGC ACCGGCTGCA GCACCCTGTG 540  
CTGCTGTCCC GCCTGGTGGT CCTGGACCAC CCCCACTACT GCAACGTCAC CTACAACGTC 600  
AATAATGGCA TCCAGTATGT GGCCAGCCAG GCGGAGCAGA ATGCGTCGGA AGTAGGCTCC 660  
CCACCCTCCT ACTCCGAGGC CTTGCTGGAC CAGAGGCCTG CGTGGTATGA CCTTCCTCCA 720  
CCGCCCTACT CTTCTGACAC GGAATCTCTG AACCAAGCCG ACCTGCCCCC CTACCGCTCC 780  
CGGTCGGGGA GTGCCAACAG TGCCAGCTCC CAGGCAGCCA PPSYSEALLD QRPAYDLPP 840  
GACACCAGCC ACAGCCCGGG GCAGCCTGGC CCCAGGAGG GCACTGCTGA GCCCAGGGAC 900  
TCTGAGCCCA GCCAGGGCAC TGAAGAAGTA TAA

Seq ID NO: 679 Protein sequence  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
| | | | | |  
MCSNGRCIPG AWQCDGLPDC FDKSDEKECP KAKSKCGPTF FPCASGIHCI IGRFRCNGFE 60  
DCPDGSDEEN CTANPLLCSY ARYHCKNGLC IDKSFICDGQ NNCQDNSDEE SCESSQEPGS 120  
GQVFVTSNQ LVYYPSTIYA IIGSSVIFVL VVALLALVLH HQKRNNLMT LPVHRLQHPV 180  
LLSRLVLDH PHHCNVYTNV NNGIQYVASQ AEQNASEVGS PPSYSEALLD QRPAYDLPP 240  
PPYSSDTESL NQADLPPYRS RSGSANSASS QAASSLLSVE DTSHPGPQPG PQEGTAEPRD 300  
SEPSQGTEEV

Seq ID NO: 680 DNA sequence  
Nucleic Acid Accession #: S78203.1  
Coding sequence: 1..2190

1 11 21 31 41 51  
| | | | | |  
ATGAATCCTT TCCAGAAAAA TGAGTCCAAG GAAACTCTTT TTTCACCTGT CTCCATTGAA 60  
GAGGTACCAC CTCGACCACC TAGCCCTCCA AAGAAGCCAT CTCCGACAAT CTGTGGCTCC 120  
AACTATCCAC TGAGCATTGC CTTCAATGTG GTGAATGAAT TCTGCGAGCG CTTTTCCTAT 180  
TATGGAATGA AAGCTGTGCT GATCCTGTAT TTCTGTATT TCCCTGCACTG GAATGAAGAT 240  
ACCTCCACAT CTATATACCA TGCCCTCAGC AGCCTCTGTT ATTTTACTCC CATCTGGGA 300  
GCAGCCATTG CTGACTCGTG GTTGGGAAAA TTCAAGACAA TCATCTATCT CTCTTGGTG 360  
TATGTGCTTG GCCATGTGAT CAAGTCCTTG GGTGCCTTAC CAATACTGGG AGGACAAGTG 420  
GTACACACAG TCCTATCATT GATCGGCTTG AGTCTAATAG CTTTGGGGAC AGGAGGCATC 480  
AAACCTGTG TGGCAGCTTT TGGTGGAGAC CAGTTTGAAG AAAACATGC AGAGGAACGG 540  
ACTAGATACT TCTCAGTCTT CTACCTGTCC ATCAATGCAG GGAGCTTGAT TTCTACATTT 600  
ATCACACCCA TGCTGAGAGG AGATGTGCAA TGTTTTGGAG AAGACTGCTA TGCATTGGCT 660  
TTTGGAGTTC CAGGACTGCT CATGGTAATT GCACTTGTG TGTTTGCAAT GGGAGCAAAA 720  
ATATACAATA AACCACCCCC TGAAGGAAAC ATAGTGCTC AAGTTTTCAA ATGTATCTGG 780  
TTTGCTATTT CCAATCGTTT CAAGAACCGT TCTGGAGACA TTCAAAGCG ACAGCACTGG 840  
CTAGACTGGG CAGCTGAGAA ATATCCAAAG CAGCTCATTA TGGATGTAAA GGCACGTACC 900  
AGGGTACTAT TCCTTTATAT CCCATTGCCG ATGTTCTGGG CTCTTTTGGG TCAGCAGGGT 960  
TCACGATGGA CTTTGCAAGC CATCAGGATG AATAGGAATT TGGGGTTTTT TGTGCTTCAG 1020  
CCGGACCAGA TGCAGTTTCT AAATCCCTTT CTGGTTCTTA TCTTCATCCC GTTGTGTTGAC 1080  
TTTGTCAATT ATCGTCTGGT CTCCAAGTGT GGAATTAAT TCTCATCACT TAGGAAAATG 1140  
GCTGTTGGTA TGATCCTAGC GTGCCTGGCA TTTGCAGTTG CGGCAGCTGT AGAGATAAAA 1200  
ATAAATGAAA TGGCCCCAGC CCAGTCAGGT CCCAGGAGG TTTTCTTACA AGTCTTGAAT 1260  
CTGGCAGATG ATGAGGTGAA GGTGACAGTG GTGGGAAATG AAAACAATTG TCTGTTGATA 1320  
GAGTCCATCA AATCCTTTCA GAAAACACCA CACTATTCCA AACTGCACCT GAAAACAAAA 1380  
AGCCAGGATT TTCATTCCA CCTGAAATAT CACAATTTGT CTCTCTACAC TGAGCATTCT 1440  
GTGCAGGAGA AGAACTGGTA CAGTCTTGTC ATTCTGTAAG ATGGGAACAG TATCTCCAGC 1500  
ATGATGGTAA AGGATACAGA AAGCAAAACA ACCAATGGGA TGACAACCGT GAGGTTTGTT 1560  
AACACTTTGC ATAAAGATGT CAACATCTCC CTGAGTACAG ATACCTCTCT CAATGTTGGT 1620  
GAAGACTATG GTGTGCTGTC TTATAGAAGT GTGCAAGAG GAGAATACCC TGCAGTGCAC 1680  
TGTAAGACAG AAGATAAGAA CTTTCTCTCT AATTTGGGTC TTCTAGACTT TGGTGCAGCA 1740  
TATCTGTTTG TTATTACTAA TAACACCAAT CAGGGTCTTC AGGCCTGGAA GATTGAAGAC 1800  
ATTCCAGCCA ACAAATGTC CATTGCGTGG CAGCTACCAC AATATGCCCT GGTACAGCT 1860  
GGGGAGGTCA TGTCTCTGT CACAGGTCTT GAGTTTTCTT ATTCTCAGGC TCCCTCTAGC 1920  
ATGAAATCTG TGCTCCAGGC AGCTTGGCTA TTGACAATTG CAGTTGGGAA TATCATCTGT 1980  
CTTGTGTTGG CACAGTTCAG TGGCCTGGTA CAGTGGGCCG AATTCATTTT GTTTTCTGTC 2040  
CTCCTGCTGG TGATCTGCCT GATCTTCTCC ATCATGGGCT ACTACTATGT TCCTGTAAAG 2100  
ACAGAGGATA TGCGGGGTCC AGCAGATAAG CACATTCTCT ACATCCAGGG GAACATGATC 2160  
AAACTAGAGA CCAAGAAGAC AAAACTCTGA

Seq ID NO: 681 Protein sequence  
Protein Accession #: AAB34388.1



1 11 21 31 41 51  
| | | | |  
MNPFOKNESK ETLFSPVSIE EVPPRPPSP KKPSPITICGS NYPLSIAFIV VNEFCERFSY 60  
YGMKAVLILY FLYFLHWNED TSTSIYHAFS SLCYFTPILG AAIADSWLKG FKTIYLSLV 120  
5 YVLGHVIVSL GALPILGGQV VHTVLSLIGL SLIALGTGGI KPCVAAFAGD QFEEKHAEER 180  
TRYFSVFYLS INAGSLISTF ITPMLRGDVQ CFGEDCYALA FGVPGLLMVI ALVVVFAMGSK 240  
IYNKPPPEGN IVAQVFKCIW FAISNRPFKNR SGDIPKRQHW LDWAAEKYPK QLIMDVKALT 300  
RVLFYIPLP MFWALLDQQG SRWTLQAIRN NRNLGFFVLQ PDQMQLNPF LVLIFIPFLD 360  
10 FVIYRLVSKC GINFSSLRKM AVGMILACLA FAVAAAVEIK INEMAPAQSG PQEVFLQVLN 420  
LADDEVKVTV VGNENNSLLI ESIKSFQKTP HYSKLHLKTK SQDFHFLKY HNLSLYTEHS 480  
VQEKWYSLV IREDGNSISS MMVKDTESKT TNGMTTVRFV NTLHKDVNIS LSTDTSLNVG 540  
EDYGVSAVRT VQRGEYPAVH CRTEDKNFSL NLGLLDFGAA YLFVITNNTN QGLQAWKIED 600  
IPANKMSIAW QLPQYALVTA GEVMSFVTGL EFSYSQAPSS MKSVLQAAWL LTIAGVNIIV 660  
15 LVVAQFSGLV QWAEFILFSC LLLVICLIFS IMGYVVPVK TEDMRGPADK HIPHIQGNMI 720  
KLETKKTKL

Seq ID NO: 682 DNA sequence  
Nucleic Acid Accession #: NM\_016077.1  
Coding sequence: 128..667

1 11 21 31 41 51  
| | | | |  
TCGCTTTGTG ATTCTTGATC CGGAACCTTG TCACCCAGGA ACCCCGGAAG AGGTAGCTCA 60  
CGCGATAGAA ACGTGTTCGC TTGCCAGAA GAAGGGAAGG CGCGAGTGAG GAAAGGAGGT 120  
25 ACTGTAGATG CCTTCCAAAT CCTTGTTAT GGAATATTTG GCTCATCCCA GTACACTCGG 180  
CTTGCTGTT GGAGTTGCTT GTGGCATGTG CCTGGGCTGG AGCCTTCGAG TATGCTTTGG 240  
GATGCTCCCC AAAAGCAAGA CGAGCAAGAT ACACACAGAT ACTGAAAGTG AAGCAAGCAT 300  
CTTGGGAGAC AGCGGGGAGT ACAAGATGAT TCTTGTGGTT CGAAATGACT TAAAGATGGG 360  
30 AAAAGGGAAA GTGGCTGCCC AGTGCTCTCA TGCTGCTGTT TCAGCCTACA AGCAGATTCA 420  
AAGAAGAAAT CCTGAAATGC TCAAACAATG GGAATACTGT GGCCAGCCCA AGGTGGTGGT 480  
CAAAGCTCCT GATGAAGAAA CCCTGATTGC ATTATTGGCC CATGCAAAAA TGCTGGGACT 540  
GACTGTAAGT TTAATTCAAG ATGCTGGACG TACTCAGATT GCACCAGGCT CTCAAACCTGT 600  
CCTAGGGATT GGGCCAGGAC CAGCAGACCT AATTGACAAA GTCAGTGGTC ACCTAAAACT 660  
35 TTAAGGTTG GACTTTGATA TGACAAACAC CCCTCCATCA CAAGTGTGTTG AAGCCTGTCA 720  
GATTCTAACA ACAAAGCTG AATTTCTTCA CCAACTTAA ATGTTCTTGA GATGAAAATA 780  
AAACCTATTC CCATGTTCTA AAAAAA

Seq ID NO: 683 Protein sequence  
Protein Accession #: NP\_057161.1

1 11 21 31 41 51  
| | | | |  
MPSKSLVMEY LAHPSTLGLA VGVACGMCLG WSLRVCFGML PKSKTSKTHT DTESEASILG 60  
DSGEYKMLV VRNDLKMKGK KVAAQCSHAA VSAYKQIQRR NPEMLKQWEY CGQPKVVVKA 120  
45 PDEETLIALL AHAKMLGLTV SLIQDAGRTQ IAPGSQTVLG IGP GPADLID KVTGHLKLY

Seq ID NO: 684 DNA sequence  
Nucleic Acid Accession #: NM\_004864.1  
Coding sequence: 26..952

1 11 21 31 41 51  
| | | | |  
CGGAACGAGG GCAACCTGCA CAGCCATGCC CGGGCAAGAA CTCAGGACGG TGAATGGCTC 60  
TCAGATGCTC CTGGTGTTCG TGGTGCTCTC GTGGCTGCCG CATGGGGGCG CCCTGTCTCT 120  
55 GGCCGAGGCG AGCCGCGCAA GTTTCCTGGG ACCCTCAGAG TTGCACTCCG AAGACTCCAG 180  
ATTCCGAGAG TTGCGGAAAC GCTACGAGGA CCTGCTAACC AGGCTGCGGG CCAACCAGAG 240  
CTGGGAAGAT TCGAACACCG ACCTCGTCCC GGCCCTGCA GTCCGGATAC TCACGCCAGA 300  
AGTGCGGCTG GGATCCGGCG GCCACCTGCA CCTGCGTATC TCTCGGGCCG CCCTTCCCGA 360  
60 GGGGCTCCCC GAGGCCTCCC GCCTTACCG GCGCTGTGTT CCGCTGTCCC CGACGGCGTC 420  
AAGGTCGTGG GACGTGACAC GACCGCTGCG GCGTCAGCTC AGCCTTGCAA GACCCCAAGC 480  
GCCCCGCGTG CACCTGCGAC TGTCGCCGCC GCCGTGCGAG TCGGACCAAC TGCTGGCAGA 540  
ATCTTCGTCC GCACGGCCCC AGCTGGAGTT GCACTTGCGG CCGCAAGCCG CCAGGGGGCG 600  
CCGCAAGAGC GTGCGCGCA ACGGGGACGA CTGTCCGCTC GGGCCCGGGC GTTGTGCGG 660  
TCTGCACACG GTCCGCGCGT CGCTGGAAGA CCTGGGCTGG GCCGATTGGG TGCTGTGCGC 720  
65 ACGGAGGTG CAAGTGACCA TGTGCATCGG CCGGTGCCCC AGCCAGTTCC GGGCGGCAAA 780  
CATGCACGCG CAGATCAAGA CGAGCCTGCA CCGCCTGAAG CCCGACACGG AGCCAGCGCC 840  
CTGCTGCGTG CCGCCAGCT ACAATCCCAT GGTGCTCATT CAAAAGACCG ACACCGGGGT 900  
GTCGCTCCAG ACCTATGATG ACTTGTTAGC CAAAGACTGC CACTGCATAT GAGCAGTCC 960  
70 GTTCCTTCCA CTGTGCACCT GCGCGGGGGA GGCGACCTCA GTTGTCTTGC CCTGTGGAAT 1020  
GGGCTCAAGG TTCCTGAGAC ACCCGATTCC TGCCCAACA GCTGTATTTA TATAAGTCTG 1080  
TTATTTATTA TTAATTTATT GGGGTGACCT TCTGGGGGAC TCGGGGGCTG GTCTGATGGA 1140  
ACTGTGTATT TATTTAAAC TCTGGTGATA AAAATAAAGC TGTCTGAAC GTTAAAAAAA 1200  
AAAA

Seq ID NO: 685 Protein sequence  
Protein Accession #: NP\_004855.1

1 11 21 31 41 51  
| | | | |  
MPGQELRTVN GSQMLLVLLV LSWLPHGGAL SLAEASRAS FPGPSELHSED SRFRELKRY 60  
EDLLTRLRAN QSWEDSNTDL VPAPAVRILT PEVRLGSGGH LHLRISRAAL PEGLPASRL 120  
80 HRLFRLSPT ASRSWDVTRP LRRQLSLARP QAPALHLRLS PPPSQSDQLL AESSSARPQL 180  
ELHLRPQAR GRRRARARNG DDCPLGPGR CRLHTVRASL EDLGWADWVL SPREVQVTMC 240  
85 IGACPSQFRA ANMHAQIKTS LHRLKPDTEP APCCVPASYN PMVLIQKTDV GVSLLQTYDDL 300  
LAKDCHCI

Seq ID NO: 686 DNA sequence



1	11	21	31	41	51	
ACCAAATCAA	CCATAGGTCC	AAGAAACAATT	GTCTCTGGAC	GGCAGCTATG	CGACTCACCG	60
TGCTGTGTGC	TGTGTGCCTG	CTGGCTGGCA	GCCTGGCCCT	GCCGCTGCCCT	CAGGAGGCCG	120
GAGGCATGAG	TGAGCTACAG	TGGGCAACAGG	CTCAGGACTA	TCTCAAGAGA	TTTATCTCT	180
ATGACTCAGA	AACAAAAAAT	GCCAACAGTT	TAGAAGCCAA	ACTCAAGGAG	ATGCAAAAAT	240
TCTTTGGCCT	ACCTATAACT	GGAATGTTAA	ACTCCGCGT	CATAGAAATA	ATGCAGAAGC	300
CCAGATGTGG	AGTGCCAGAT	GTTGCAGAAT	ACTCACTATT	TCCAAATAGC	CCAAAATGGA	360
CTTCCAAAGT	GGTCACCTAC	AGGATCGTAT	CATATACTCG	AGACTTACCG	CATATTACAG	420
TGGATCGATT	AGTGTCAAAG	GCTTTAAACA	TGTGGGGCAA	AGAGATCCCC	CTGCATTTCA	480
GGAAAGTTGT	ATGGGGAACT	GCTGACATCA	TGATTGGCTT	TGCGCGAGGA	GCTCATGGGG	540
ATCTCTACCC	ATTTGATGGG	CCAGGAAACA	CGCTGGCTCA	TGCCTTTGCG	CCTGGGACAG	600
GTCTCGGAGG	AGATGCTCAC	TTCGATGAGG	ATGAACGCTG	GACGGATGGT	AGCAGTCTAG	660
GGATTAACTT	CCTGTATGCT	GCAACTCATG	AACTTGGCCA	TTCTTTGGGT	ATGGGACATT	720
CCTCTGATCC	TAATGCAGTG	ATGTATCCAA	CCTATGGAAA	TGGAGATCCC	CAAAATTTTA	780
AACTTTCCCA	GGATGATATT	AAAGGCATTC	AGAACTATA	TGGAAAGAGA	AGTAATTCAA	840
GAAAGAAATA	GAAACTTCAG	GCAGAACATC	CATTCAATTA	TTCATTGGAT	TGTATATCAT	900
TGTTGCACAA	TCAGAAATTA	TAAGCACTGT	TCCTCCACTC	CATTTAGCAA	TTATGTCAAC	960
CTTTTTTATT	GCAGTTGGTT	TTTGAATGTC	TTTCACTCCT	TTTATTGGTT	AAACTCCTTT	1020
ATGTTGTGAC	TGTGTCTTAT	TCCATCTATG	AGCTTTGTCA	GTGCGCTAG	ATGTCAATAA	1080
ATGTTACATA	CACAAATAAA	TAAAATGTTT	ATTCCATGGT	AAATTTA		

1	11	21	31	41	51	
MRLTVLCAVC	LLPGSLALPL	PQEAGGMSEL	QWEQAQDYLK	RFYLYDSETK	NANSLEAKLK	60
EMQKFFGLPI	TGMLNSRVIE	IMQKPRCGVP	DVAEYSLFPN	SPKWTSKVVV	YRIVSYTRDL	120
PHITVDRLS	KALNMWGKEI	PLHFRKVVWG	TADIMIGFAR	GAHGDSYPFD	GPGNTLAHAF	180
APGTGLGSDA	HFDEDERWTD	QSSGLGINFLY	AATHELGHSL	GMGHSSDPNA	VMYPTYGNGD	240
PQFKLSQDD	IKGIQKLYGK	RSNSRKK				

1	11	21	31	41	51	
ATGACAGGAG	TGTTTGACAG	AAGGGTCCCC	AGCATCCGAT	CCGGCGACTT	CCAAGCTCCG	60
TTCCAGACGT	CCGCAGCTAT	GCACCATCCG	TCTCAGGAAT	CGCCAACCTT	GCCCCGAGTCT	120
TCAGCTACCG	ATTCTGACTA	CTACAGCCCT	ACGGGGGGAG	CCCCGCACGG	CTACTGCTCT	180
CCTACCTCGG	CTTCTATGG	CAAAGCTCTC	AACCCTTACC	AGTATCAGTA	TCACGGCGTG	240
AACGGCTCCG	CCGGGAGCTA	CCCAGCCAAA	GCTTATGCCG	ACTATAGCTA	CGCTAGCTCC	300
TACCACCACT	ACGGCGGGCG	CTACAACCGC	GTCCCAAGCG	CCACCAACCA	GCCAGAGAAA	360
GAAGTGACCG	AGCCCCAGGT	GAGAATGGTG	AATGGCAAAC	CAAAGAAAGT	TCGTAAACCC	420
AGGACTATTT	ATTCCAGCTT	TCAGCTGGCC	GCATTACAGA	GAAGGTTTCA	GAAGACTCAG	480
TACCTCGCCT	TGCCGGAACG	CGCCGAGCTG	GCCGCCTCGC	TGGGATTGAC	ACAAACACAG	540
GTGAAAATCT	GGTTCAGAA	CAAAAGATCC	AAGATCAAGA	AGATCATGAA	AAACGGGGAG	600
ATGCCCCCGG	AGCACAGTCC	CAGTCCAGC	GACCCAATGG	CGTGTAACTC	GCCGAGTCT	660
CCAGCGGTGT	GGGAGCCCCA	GGGCTCGTCC	CGCTCGCTCA	GCCACCACCC	TCATGCCCCA	720
CCTCCGACCT	CCAACCAATC	CCAGCGTCC	AGCTACCTGG	AGAACTCTGC	ATCCTGGGTAC	780
ACAAGTGCAG	CCAGCTCAAT	CAATTCCAC	CTGCCGCCGC	CGGGCTCCTT	ACAGCACCCG	840
CTGGCGCTGG	CCTCCGGGAC	ACTCTATTAG				

1	11	21	31	41	51	
MTGVFDRRVP	SIRSGDFQAP	FQTSAAMHHP	SQESPTLPES	SATDSDYVSP	TGGAPHGYCS	60
PTSASYGKAL	NPYQYQYHGV	NGSAGSYPAK	AYADYSYASS	YHQYGGAYNR	VPSATNQPEK	120
EVTEPEVRMV	NGKPKKVRKP	RTIYSSFQLA	ALQRRFQKTQ	YLALPERAEL	AASLGLTQTQ	180
VKIWFQNKRS	KIKKIMKNGE	MPPEHSPSSS	DPMACNSPQS	PAVWEPQGSS	RSLSHHPPAH	240
PPTSNSQSPAS	SYLENSASWY	TSAASSINSH	LPPPGSLQHP	LALASGTLV		



It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.



WHAT IS CLAIMED IS:

- 1                   1.       A method of detecting a lung cancer-associated transcript in a cell  
2       from a patient, the method comprising contacting a biological sample from the patient with a  
3       polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence  
4       as shown in Tables 1A-16.
- 1                   2.       The method of claim 1, wherein the polynucleotide selectively  
2       hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16.
- 1                   3.       The method of claim 1, wherein the biological sample is a tissue  
2       sample.
- 1                   4.       The method of claim 1, wherein the biological sample comprises  
2       isolated nucleic acids.
- 1                   5.       The method of claim 4, wherein the nucleic acids are mRNA.
- 1                   6.       The method of claim 4, further comprising the step of amplifying  
2       nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1                   7.       The method of claim 1, wherein the polynucleotide comprises a  
2       sequence as shown in Tables 1A-16.
- 1                   8.       The method of claim 1, wherein the polynucleotide is labeled.
- 1                   9.       The method of claim 8, wherein the label is a fluorescent label.
- 1                   10.      The method of claim 1, wherein the polynucleotide is immobilized on  
2       a solid surface.
- 1                   11.      The method of claim 1, wherein the patient is undergoing a therapeutic  
2       regimen to treat lung cancer.
- 1                   12.      The method of claim 1, wherein the patient is suspected of having lung  
2       cancer.
- 1                   13.      A method of monitoring the efficacy of a therapeutic treatment of lung  
2       cancer, the method comprising the steps of:



3 (i) providing a biological sample from a patient undergoing the therapeutic  
4 treatment; and

5 (ii) determining the level of a lung cancer-associated transcript in the  
6 biological sample by contacting the biological sample with a polynucleotide that selectively  
7 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16,  
8 thereby monitoring the efficacy of the therapy.

1 14. The method of claim 13, further comprising the step of: (iii) comparing  
2 the level of the lung cancer-associated transcript to a level of the lung cancer-associated  
3 transcript in a biological sample from the patient prior to, or earlier in, the therapeutic  
4 treatment.

1 15. The method of claim 13, wherein the patient is a human.

1 16. A method of monitoring the efficacy of a therapeutic treatment of lung  
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic  
4 treatment; and

5 (ii) determining the level of a lung cancer-associated antibody in the biological  
6 sample by contacting the biological sample with a polypeptide encoded by a polynucleotide  
7 that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in  
8 Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated  
9 antibody, thereby monitoring the efficacy of the therapy.

1 17. The method of claim 16, further comprising the step of: (iii) comparing  
2 the level of the lung cancer-associated antibody to a level of the lung cancer-associated  
3 antibody in a biological sample from the patient prior to, or earlier in, the therapeutic  
4 treatment.

1 18. The method of claim 16, wherein the patient is a human.

1 19. A method of monitoring the efficacy of a therapeutic treatment of lung  
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic  
4 treatment; and



5 (ii) determining the level of a lung cancer-associated polypeptide in the  
6 biological sample by contacting the biological sample with an antibody, wherein the antibody  
7 specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to  
8 a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby  
9 monitoring the efficacy of the therapy.

1 20. The method of claim 19, further comprising the step of: (iii) comparing  
2 the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated  
3 polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic  
4 treatment.

1 21. The method of claim 19, wherein the patient is a human.

1 22. An isolated nucleic acid molecule consisting of a polynucleotide  
2 sequence as shown in Tables 1A-16.

1 23. The nucleic acid molecule of claim 22, which is labeled.

1 24. The nucleic acid of claim 23, wherein the label is a fluorescent label

1 25. An expression vector comprising the nucleic acid of claim 22.

1 26. A host cell comprising the expression vector of claim 25.

1 27. An isolated polypeptide which is encoded by a nucleic acid molecule  
2 having polynucleotide sequence as shown in Tables 1A-16.

1 28. An antibody that specifically binds a polypeptide of claim 27.

1 29. The antibody of claim 28, further conjugated to an effector component.

1 30. The antibody of claim 29, wherein the effector component is a  
2 fluorescent label.

1 31. The antibody of claim 29, wherein the effector component is a  
2 radioisotope or a cytotoxic chemical.

1 32. The antibody of claim 29, which is an antibody fragment.



- 1                    33.     The antibody of claim 29, which is a humanized antibody
- 1                    34.     A method of detecting a lung cancer cell in a biological sample from a  
2 patient, the method comprising contacting the biological sample with an antibody of claim  
3 28.
- 1                    35.     The method of claim 34, wherein the antibody is further conjugated to  
2 an effector component.
- 1                    36.     The method of claim 35, wherein the effector component is a  
2 fluorescent label.
- 1                    37.     A method of detecting antibodies specific to lung cancer in a patient,  
2 the method comprising contacting a biological sample from the patient with a polypeptide  
3 encoded by a nucleic acid comprises a sequence from Tables 1A-16.
- 1                    38.     A method for identifying a compound that modulates a lung cancer-  
2 associated polypeptide, the method comprising the steps of:  
3                    (i) contacting the compound with a lung cancer-associated polypeptide, the  
4 polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least  
5 80% identical to a sequence as shown in Tables 1A-16; and  
6                    (ii) determining the functional effect of the compound upon the polypeptide.
- 1                    39.     The method of claim 38, wherein the functional effect is a physical  
2 effect.
- 1                    40.     The method of claim 38, wherein the functional effect is a chemical  
2 effect.
- 1                    41.     The method of claim 38, wherein the polypeptide is expressed in a  
2 eukaryotic host cell or cell membrane.
- 1                    42.     The method of claim 38, wherein the functional effect is determined by  
2 measuring ligand binding to the polypeptide.
- 1                    43.     The method of claim 38, wherein the polypeptide is recombinant.



1           44.     A method of inhibiting proliferation of a lung cancer-associated cell to  
2     treat lung cancer in a patient, the method comprising the step of administering to the subject a  
3     therapeutically effective amount of a compound identified using the method of claim 38.

1           45.     The method of claim 44, wherein the compound is an antibody.

1           46.     The method of claim 45, wherein the patient is a human.

1           47.     A drug screening assay comprising the steps of  
2                 (i) administering a test compound to a mammal having lung cancer or a cell  
3     isolated therefrom;  
4                 (ii) comparing the level of gene expression of a polynucleotide that selectively  
5     hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16 in a  
6     treated cell or mammal with the level of gene expression of the polynucleotide in a control  
7     cell or mammal, wherein a test compound that modulates the level of expression of the  
8     polynucleotide is a candidate for the treatment of lung cancer.

1           48.     The assay of claim 47, wherein the control is a mammal with lung  
2     cancer or a cell therefrom that has not been treated with the test compound.

1           49.     The assay of claim 47, wherein the control is a normal cell or mammal.

1           50.     A method for treating a mammal having lung cancer comprising  
2     administering a compound identified by the assay of claim 47.

1           51.     A pharmaceutical composition for treating a mammal having lung  
2     cancer, the composition comprising a compound identified by the assay of claim 47 and a  
3     physiologically acceptable excipient.



(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number  
**WO 02/086443 A2**

(51) International Patent Classification<sup>7</sup>: **G01N**

(21) International Application Number: PCT/US02/12476

(22) International Filing Date: 18 April 2002 (18.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/284,770	18 April 2001 (18.04.2001)	US
60/290,492	10 May 2001 (10.05.2001)	US
60/339,245	9 November 2001 (09.11.2001)	US
60/350,666	13 November 2001 (13.11.2001)	US
60/334,370	29 November 2001 (29.11.2001)	US
60/372,246	12 April 2002 (12.04.2002)	US

(71) Applicant (for all designated States except US): **EOS BIOTECHNOLOGY, INC.** [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **AZIZ, Natasha** [US/US]; 411 California Avenue, Palo Alto, CA 94306 (US). **MURRAY, Richard** [US/US]; 22643 Woodbridge Court, Cupertino, CA 95014 (US).

(74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

— of inventorship (Rule 4.17(iv)) for US only

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

(57) Abstract: Described herein are methods and compositions that can be used for diagnosis and treatment of lung cancer and similar pathologies. Also described herein are methods that can be used to identify modulators of lung cancer and similar pathologies.



**WO 02/086443 A2**



# METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

5

## CROSS-REFERENCES TO RELATED APPLICATIONS

This application is related to USSN 60/284,770, filed April 18, 2001; USSN 60/290,492, filed May 10, 2001; USSN 60/334,370, filed November 29, 2001; USSN 60/339,245, filed November 9, 2001; USSN 60/350,666, filed November 13, 2001; and  
10 USSN 60/xxx,xxx, filed April 12, 2002 (Docket OMNI-002P); each of which is incorporated herein by reference in its entirety.

## FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression  
15 profiles and nucleic acids, products, and antibodies thereto that are involved in lung cancer; and to the use of such expression profiles and compositions in diagnosis and therapy of lung cancer. The invention further relates to methods for identifying and using agents and/or targets that inhibit lung cancer or related conditions.

20

## BACKGROUND OF THE INVENTION

Lung cancer is the second most commonly occurring cancer in the United States and is the leading cause of cancer-related death. It is estimated that there are over 160,000 new cases of lung cancer in the United States every year. Of those who are diagnosed with lung cancer, 86 percent will die within five years. Lung cancer is the most common visceral  
25 cancer in men and accounts for nearly one third of all cancer deaths in both men and women. In fact, lung cancer accounts for 7% of all deaths, due to any cause, in both men and women.

Smoking is the primary cause of lung cancer, with more than 80% of lung cancers resulting from smoking. About 400 to 500 separate gaseous substances are present in the smoke of a non-filter cigarette. The most noteworthy substances include nitrogen oxides,  
30 hydrogen cyanide, formaldehyde, benzene, and toluene. The particles present in cigarette smoke contain at least 3,500 individual compounds such as nicotine, tobacco alkaloids (nornicotine, anatabine, anabasine), polycyclic aromatic hydrocarbons (e.g., benzo(a)pyrene, B(a)P), naphthalenes, aromatic amines, phenols, and tobacco-specific nitrosamines.



Tobacco-specific nitrosamines are formed during tobacco curing and processing, and are suspected of causing lung cancer in humans. In rodent studies, regardless of the where or how it is applied, the tobacco-specific nitrosamine known as NNK produces lung adenomas and lung adenocarcinomas. The tobacco-specific nitrosamine known as NNAL also produces  
5 lung adenocarcinomas in rodents.

Many of the chemicals found in cigarette smoke also affect the nonsmoker inhaling "secondhand" or sidestream smoke. Indeed, the smoke inhaled by non-smokers has a chemical composition similar to the smoke inhaled by smokers, but, importantly, the concentrations of the carcinogenic tobacco-specific nitrosamines are present in higher  
10 concentrations in second hand smoke. For this and other reasons, "passive smoking" is an important cause of lung cancer, causing as many as 3,000 lung cancer deaths in nonsmokers each year.

In addition to smoking, other factors thought to be causes of lung cancer include on-the-job exposure to carcinogens such as asbestos and uranium, exposure to chemical hazards  
15 such as radon, polycyclic aromatic hydrocarbons, chromium, nickel, and inorganic arsenic, genetic factors, and diet.

Histological classification of various lung cancers define the types of cancer that begin in the lung. See, e.g., Travis, et al. (1999) Histological Typing of Lung and Pleural Tumours (International Histological Classification of Tumours, No 1. Four major cell types  
20 make up more than 88% of all primary lung neoplasms. These are: squamous or epidermoid carcinoma, small cell (also called oat cell) carcinoma, adenocarcinoma, and large cell (also called large cell anaplastic) carcinoma. The remainder include undifferentiated carcinomas, carcinoids, bronchial gland tumors, and other rarer types. The various cell types have different natural histories and responses to therapy, and, thus, a correct histologic diagnosis is  
25 the first step of effective treatment.

Small cell lung cancer (SCLC) accounts for 18-25% of all lung cancers, and occurs less frequently than non-small cell lung cancers, and generally spread to distant organs more rapidly than non-small cell lung cancer. In general, at the time of presentation small cell lung cancers have already spread beyond the bounds where surgery and curative intent  
30 can be undertaken. However, if identified early enough, these cancers are often responsive to chemotherapy and thoracic radiation treatment.

Non-small cell lung cancers (NSCLC) are the more frequently occurring form of lung cancer. They comprise squamous cell carcinoma, adenocarcinoma, and large cell carcinoma



and account for more than 75% of all lung cancers. Non-small cell tumors that are localized at the time of presentation can sometimes be cured with surgery and/or radiotherapy, but usually are not identified until significant metastasis has occurred, which are typically not very responsive to surgical, chemotherapy, or radiation treatment..

5 The screening of asymptomatic persons at high risk for lung cancer has often proven ineffective. In general, only 5 to 15 percent of lung cancer patients have their disease detected while they are asymptomatic. Of course, early detection and treatment are critical factors in the fight against lung cancer. The average survival rate is 49% for those whose cancer is detected early, before the cancer has spread from the lung. Lung cancer often  
10 spreads outside of the lung, and it may have spread to the bones or brain by the time it is diagnosed. While the prognosis may be better for lung cancers that are detected early, because of the lack ofv effective curative treatments, early detection does not necessarily alter the total death rate from lung cancer.

Thus, methods for diagnosis and prognosis of lung cancer and effective treatment of  
15 lung cancer would be desirable. Accordingly, provided herein are methods that can be used in diagnosis and prognosis of lung cancer. Further provided are methods that can be used to screen candidate therapeutic agents for the ability to modulate, e.g., treat, lung cancer. Additionally, provided herein are molecular targets and compositions for therapeutic intervention in lung disease and other metastatic cancers.

## 20 SUMMARY OF THE INVENTION

The present invention provides nucleotide sequences of genes that are up- and down-regulated in lung cancer cells. Such genes are useful for diagnostic purposes, and also as targets for screening for therapeutic compounds that modulate lung cancer, such as  
25 antibodies. The methods of detecting nucleic acids of the invention or their encoded proteins can be used for a number of purposes. Examples include early detection of lung cancers, monitoring and early detection of relapse following treatment of lung cancers, monitoring response to therapy of lung cancers, determining prognosis of lung cancers, directing therapy of lung cancers, selecting patients for postoperative chemotherapy or radiation therapy,  
30 selecting therapy, determining tumor prognosis, treatment, or response to treatment, and early detection of precancerous lesions of the lung. Examples of benign or precancerous lesions include: atelectasis, emphysema, brochitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis (HP), interstitial pulmonary fibrosis (IPF), asthma, and



bronchiectasis. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting a lung cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16. Alternatively, the sample may be contacted with a specific binding reagent, e.g., antibody.

In one embodiment, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16. In another embodiment, the polynucleotide comprises a sequence as shown in Tables 1A-16.

In one embodiment, the biological sample is a tissue sample, or a body fluid. In another embodiment, the biological sample comprises isolated nucleic acids, e.g., mRNA.

In one embodiment, the polynucleotide is labeled, e.g., with a fluorescent label. In one embodiment, the polynucleotide is immobilized on a solid surface. In one embodiment, the patient is undergoing a therapeutic regimen to treat lung cancer. In another embodiment, the patient is suspected of having lung cancer. In one embodiment, the patient is a primate, e.g., a human.

In one embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy. Or the sample may be evaluated for protein, e.g., contacting the sample with an antibody.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated transcript to a level of the lung cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. Or the sample may be evaluated for comparison of protein.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a



biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated antibody in the biological sample by contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated antibody, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated antibody to a level of the lung cancer-associated antibody in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated polypeptide in the biological sample by contacting the biological sample with an antibody, wherein the antibody specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1A-16. In one embodiment, an expression vector or cell comprises the isolated nucleic acid. In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16. In one embodiment, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical. In one embodiment, the antibody is an antibody fragment. In another embodiment, the antibody is humanized.

In one aspect, the present invention provides a method of detecting lung cancer in a patient, the method comprising contacting a biological sample from the patient with an antibody or protein as described herein.



In another aspect, the present invention provides a method of detecting antibodies specific to a lung cancer gene in a patient, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprises a sequence from Tables 1A-16.

5 In another aspect, the present invention provides a method for identifying a compound that modulates a lung cancer-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a lung cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16; and (ii) determining the functional effect of the  
10 compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect. In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant. In one embodiment, the functional effect is determined by measuring ligand binding to the  
15 polypeptide.

In another aspect, the present invention provides a method of inhibiting proliferation or another critical process of a lung cancer-associated cell to treat lung cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein. In one embodiment, the compound is an  
20 antibody.

In another aspect, the present invention provides a drug screening assay comprising the steps of: (i) administering a test compound to a mammal having lung cancer or a cell isolated therefrom; (ii) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables  
25 1A-16 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of lung cancer.

In one embodiment, the control is a mammal with lung cancer or a cell therefrom that has not been treated with the test compound. In another embodiment, the control is a normal  
30 cell or mammal, or a non-malignant lung disease.

In another aspect, the present invention provides a method for treating a mammal having lung cancer comprising administering a compound identified by the assay described herein.



In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having lung cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

5

## DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and treatment of lung disease or cancer, as well as methods for screening for compositions which modulate lung cancer. "Treatment, monitoring, detection or modulation of lung disease or cancer" includes treatment, monitoring, detection, or  
10 modulation of lung disease in those patients who have lung disease (whether malignant or non-malignant, e.g., emphysema, bronchitis, or fibrosis) as well as patients with lung cancers in which gene expression from a gene in Tables 1A-16 is increased or decreased, indicating that the subject is more likely to have disease. In particular, while these targets are identified primarily from lung cancer samples, these same targets are likely to be similarly found in  
15 analyses of other medical conditions. These other conditions may result from similar pathological processes which affect similar tissues, e.g., lung cancer, small cell lung carcinoma (oat cell carcinoma), non-small cell carcinomas (e.g., squamous cell carcinoma, adenocarcinoma, large cell lung carcinoma, carcinoid, granulomatous), fibrosis (idiopathic pulmonary fibrosis (IPF), hypersensitivity pneumonitis (HP), interstitial pneumonitis,  
20 nonspecific idiopathic pneumonitis (NSIP)), chronic obstructive pulmonary disease (COPD, e.g., emphysema, chronic bronchitis), asthma, bronchiectasis, and esophageal cancer. See, e.g., the NCI webpage and USSN 60/347,349 and USSN 60/xxx,xxx (docket LFBR-001-1P, filed March 29, 2002), each of which is incorporated herein by reference. The treatment may be of lung cancer or related condition itself, or treatment of metastasis.

25

In particular, identification of markers selectively expressed on these cancers allows for use of that expression in diagnostic, prognostic, or therapeutic methods. As such, the invention defines various compositions, e.g., nucleic acids, polypeptides, antibodies, and small molecule agonists/antagonists, which will be useful to selectively identify those markers. For example, therapeutic methods may take the form of protein therapeutics which  
30 use the marker expression for selective localization or modulation of function (for those markers which have a causative disease effect), for vaccines, identification of binding partners, or antagonism, e.g., using antisense or RNAi. The markers may be useful for molecular characterization of subsets of lung diseases, which subsets may actually require



very different treatments. Moreover, the markers may also be important in related diseases to the specific cancers, e.g., which affect similar tissues in non-malignant diseases, or have similar mechanisms of induction/maintenance. Metastatic processes or characteristics may also be targeted. Diagnostic and prognostic uses are made available, e.g., to subset related but distinct diseases, or to determine treatment strategy. The detection methods may be based upon nucleic acid, e.g., PCR or hybridization techniques, or protein, e.g., ELISA, imaging, IHC, etc. The diagnosis may be qualitative or quantitative, and may detect increases or decreases in expression levels.

Tables 1A-16 provide unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in lung cancer samples. The tables also provide an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. In Table 1A, genes marked as “target 1” or “target 2” are particularly useful as therapeutic targets. Genes marked as “target 3” are particularly useful as diagnostic markers. Genes marked as “chron” are upregulated in chronically diseased lung (e.g., emphysema, bronchitis, fibrosis) relative to lung tumors and normal tissue. In certain analyses, the ratio for the “chron” category was determined using the 70th percentile of chronically diseased lung samples divided by the 90th percentile of normal lung samples. The ratio for the targets was determined using the 70th percentile of lung tumor samples divided by the 90th percentile of normal lung samples.

## Definitions

The term “lung cancer protein” or “lung cancer polynucleotide” or “lung cancer-associated transcript” refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1A-16 and conservatively modified variants thereof; or (4)



have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence

5 encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "lung cancer polypeptide" and a "lung cancer polynucleotide," include both naturally occurring or recombinant forms.

10 A "full length" lung cancer protein or nucleic acid refers to a lung cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains the elements normally contained in one or more naturally occurring, wild type lung cancer polynucleotide or polypeptide sequences. The "full length" may be prior to, or after, various stages of post-translational processing or splicing, including alternative splicing.

15 "Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a lung cancer protein, polynucleotide, or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes, archival materials, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc.

20 Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or other mammal; or a bird; reptile; 25 fish. Livestock and domestic animals are of interest.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), or by performing the 30 methods of the invention in vivo. Archival tissues or materials, having treatment or outcome history, will be particularly useful.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the



same or have a specified percentage of amino acid residues or nucleotides that are the same (e.g., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using, e.g., a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be “substantially identical.” This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or insertions, substitutions, and naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A “comparison window”, as used herein, includes reference to a segment of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat’l. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer



Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel, et al. (eds. 1995 and supplements) Current Protocols in Molecular Biology.

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul, et al. (1977) Nuc. Acids Res. 25:3389-3402 and Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between



two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be negative large numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture Collection catalog or web site, [www.atcc.org](http://www.atcc.org)).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least about 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant or component from the composition to be purified.



In this sense, purification does not require that the purified compound be homogeneous, e.g., 100% pure.

The terms “polypeptide,” “peptide” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline,  $\gamma$ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refer to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an  $\alpha$  carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain some basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refer to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function similarly to another amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

“Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG, and GCU each encode the amino acid alanine. Thus, at each position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of



conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. In certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally similar molecule. Accordingly, a silent variation of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not necessarily with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. Typically conservative substitutions include for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, Proteins (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts, et al. (1994) Molecular Biology of the Cell (3<sup>rd</sup> ed.) and Cantor and Schimmel (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules. "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of  $\beta$ -sheet and  $\alpha$ -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.



“Nucleic acid” or “oligonucleotide” or “polynucleotide” or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are a polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have at least one different linkage, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, in Sanghui and Cook, eds. Carbohydrate Modifications in Antisense Research, ASC Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature ( $T_m$ ) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4° C drop in  $T_m$  for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9° C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary



strand; thus the sequences described herein also provide the complement of the sequence.

The nucleic acid may be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine

5 hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g., the individual units of a peptide nucleic  
10 acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, physiological, chemical, or other physical means. For example, useful labels include  $^{32}\text{P}$ , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins  
15 or other entities which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The labels may be incorporated into the cancer nucleic acids, proteins, and antibodies. Many methods known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter, et al. (1962) Nature 144:945; David, et al. (1974) Biochemistry  
20 13:1014-1021; Pain, et al. (1981) J. Immunol. Meth., 40:219-230; and Nygren (1982) J. Histochem. and Cytochem. 30:407-412.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody.  
25 The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either  
30 covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method



using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a “nucleic acid probe or oligonucleotide” is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, e.g., through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, preferably one that does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled, e.g., with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled, e.g., with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term “recombinant” when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term “recombinant nucleic acid” herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered



recombinant for the purposes of the invention. Similarly, a “recombinant protein” is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

5 The term “heterologous” when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two  
10 or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A “promoter” is typically an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type  
15 promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A “constitutive” promoter is a promoter that is active under most environmental and developmental conditions. An “inducible” promoter is a promoter that is active under environmental or developmental regulation. The term “operably linked” refers  
20 to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, e.g., wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An “expression vector” is a nucleic acid construct, generated recombinantly or  
25 synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed in operable linkage to a promoter.

The phrase “selectively (or specifically) hybridizes to” refers to the binding,  
30 duplexing, or hybridizing of a molecule selectively to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).



The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to essentially no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in "Overview of principles of hybridization and the strategy of nucleic acid assays" in Tijssen (1993) Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes (vol. 24) Elsevier. Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at  $T_m$ , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is typically at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions are often: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C. For PCR, a temperature of about 36° C is typical for low stringency amplification, although annealing temperatures may vary between about 32° C and 48° C depending on primer length. For high stringency PCR amplification, a temperature of about 62° C is typical, although high stringency annealing temperatures can range from about 50° C to about 65° C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90° C - 95° C for 0.5 - 2 min., an annealing phase lasting 0.5 - 2 min., and an extension phase of about 72° C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis, et al.(1990) PCR Protocols, A Guide to Methods and Applications.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This



occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary “moderately stringent hybridization conditions” include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C, and a wash in 1X SSC at 45° C. A positive hybridization is at least twice background. Alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., Ausubel, et al. (ed.) Current Protocols in Molecular Biology Lippincott.

The phrase “functional effects” in the context of assays for testing compounds that modulate activity of a lung cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the lung cancer protein or nucleic acid, e.g., a physiological, enzymatic, functional, physical, or chemical effect, such as the ability to decrease lung cancer. It includes ligand binding activity; cell viability, cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein expression in cells undergoing metastasis, and other characteristics of lung cancer cells. “Functional effects” include *in vitro*, *in vivo*, and *ex vivo* activities.

By “determining the functional effect” is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a lung cancer protein sequence, e.g., physiological, functional, enzymatic, physical, or chemical effects. Such functional effects can be measured by many means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the lung cancer protein; measuring binding activity or binding assays, e.g., binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on lung cancer can also be performed using lung cancer assays known to those of skill in the art such as an *in vitro* assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein



expression in cells undergoing metastasis, and other characteristics of lung cancer cells. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for lung cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase,  $\beta$ -gal, GFP, and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

“Inhibitors”, “activators”, and “modulators” of lung cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using *in vitro* and *in vivo* assays of lung cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of lung cancer proteins, e.g., antagonists. Antisense or inhibitory nucleic acids may seem to inhibit expression and subsequent function of the protein.

“Activators” are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate lung cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of lung cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the lung cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of lung cancer can also be identified by incubating lung cancer cells with the test compound and determining increases or decreases in the expression of 1 or more lung cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more lung cancer proteins, such as lung cancer proteins encoded by the sequences set out in Tables 1A-16.

Samples or assays comprising lung cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of a lung cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more



preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase “changes in cell growth” refers to any change in cell growth and proliferation characteristics *in vitro* or *in vivo*, such as cell viability, formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. See, e.g., Freshney (1994) Culture of Animal Cells a Manual of Basic Technique pp. 231-241 (3<sup>rd</sup> ed.).

“Tumor cell” refers to precancerous, cancerous, and normal cells in a tumor.

“Cancer cells,” “transformed” cells, or “transformation” in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy (see, Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3<sup>rd</sup> ed.)).

“Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD, and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul, Fundamental Immunology.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible



for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce  $F(ab)'_2$ , a dimer of Fab which itself is a light chain joined to  $V_H-C_{H1}$  by a disulfide bond. The  $F(ab)'_2$  may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the  $F(ab)'_2$  dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Paul (ed. 1999) Fundamental Immunology (4th ed.)). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty, et al. (1990) Nature 348:552-554).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler and Milstein (1975) Nature 256:495-497; Kozbor, et al. (1983) Immunology Today 4:72; Cole, et al. (1985), pp. 77-96 in Monoclonal Antibodies and Cancer Therapy; Coligan (1991 and supplements) Current Protocols in Immunology; Harlow and Lane (1988) Antibodies, A Laboratory Manual; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty, et al. (1990) Nature 348:552-554; Marks, et al. (1992) Biotechnology 10:779-783).

A "chimeric antibody" is an antibody molecule in which, e.g, (a) the constant region, or a portion thereof, is altered, replaced, or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function, and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the



variable region, or a portion thereof, is altered, replaced, or exchanged with a variable region having a different or altered antigen specificity.

### Identification of lung cancer-associated sequences

5           In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a “fingerprint” of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is  
10   characteristic of the state of the cell. That is, normal tissue may be distinguished from cancerous or metastatic cancerous tissue, or metastatic cancerous tissue can be compared with tissue from surviving cancer patients. By comparing expression profiles of tissue in known different lung cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.  
15   Molecular profiling may distinguish subtypes of a currently collective disease designation, e.g., different forms of lung cancer (chronic disease, adenocarcinoma, etc.)

          The identification of sequences that are differentially expressed in lung cancer versus non-lung cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-  
20   regulate lung cancer, and thus tumor growth or recurrence, in a particular patient. Alternatively, a treatment step may induce other markers which may be used as targets to destroy tumor cells. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Malignant disease may be compared to non-malignant conditions. Metastatic tissue can also be analyzed to determine  
25   the stage of lung cancer in the tissue, or origin of primary tumor, e.g., metastasis from a remote primary site. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the lung cancer expression profile. This may be done by making biochips comprising sets of the important lung cancer genes,  
30   which can then be used in these screens. PCR methods may be applied with selected primer pairs, and analysis may be of RNA or of genomic sequences. These methods can also be done on the protein basis; that is, protein expression levels of the lung cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the lung cancer



nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the lung cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs or as protein or DNA vaccines.

5           Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in lung cancer relative to normal tissues and/or non-malignant lung disease, or in different types of lung disease, herein termed "lung cancer sequences." As outlined below, lung cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in lung cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the lung cancer sequences are from humans; however, as will be appreciated by those in the art, lung cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other lung cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets (dogs, cats, etc.). Lung cancer sequences from other organisms may be obtained using the techniques outlined below.

          Lung cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, lung cancer nucleic acid sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the lung cancer sequences can be generated.

          A lung cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, e.g., using homology programs or hybridization conditions.

          For identifying lung cancer-associated sequences, the lung cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, cancer and non-malignant conditions, non-malignant conditions and normal tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing lung cancer samples with metastatic cancer samples from other cancers, such as, breast, other gastrointestinal cancers, prostate, ovarian,



etc. Samples of, non metastatic disease tissue and tissue undergoing metastasis are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, e.g., from Affymetrix, Santa Clara, CA. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal lung, but also including, and not limited to colon, heart, brain, liver, breast, kidney, muscle, prostate, small intestine, large intestine, spleen, bone, and/or placenta. In a preferred embodiment, those genes identified during the lung cancer screen that are expressed in significant amounts in other tissues (e.g., essential organs) are removed from the profile, although in some embodiments, this is not necessary (e.g., where organs may be dispensable at a later stage of life). That is, when screening for drugs, it is usually preferable that the target expression be disease specific, to minimize possible side effects on other organs.

In a preferred embodiment, lung cancer sequences are those that are up-regulated in lung cancer; that is, the expression of these genes is higher in cancerous tissue than in normal lung or other tissue. "Up-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater. Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. Unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, e.g., Benson, DA, et al (1998) *Nucleic Acids Research* 26:1-7 and <http://www.ncbi.nlm.nih.gov/>. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ). Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. In some situations, the sequences may be derived from assembly of available sequences or be predicted from genomic DNA using exon prediction algorithms, such as FGENESH (Salamov and Solovyev (2000) *Genome Res.* 10:516-522). In other situations, sequences have been derived from cloning and sequencing of isolated nucleic acids.

In another preferred embodiment, lung cancer sequences are those that are down-regulated in the lung cancer; that is, the expression of these genes is lower in cancerous tissue



or normal lung or other tissue. "Down-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater, or, when the ratio is presented as a number less than one, that the ratio is less than one, preferably 0.5 or less, more preferably 0.25 or less.

5

### Informatics

The ability to identify genes that are over or under expressed in lung cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with lung cancer. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (see Anderson (1998) Pharmaceutical Proteomics: Targets, Mechanism, and Function, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in a form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample representing lung cancer, i.e., the identification of lung cancer-associated sequences described herein, provide an abundance of information, which can be correlated with



pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

See also Mount, et al. (2001) Bioinformatics; Durbin, et al. (eds., 1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids (; Baxevas and Ouellette (eds., 1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins); Rashidi and Buehler (1999) Bioinformatics: Basic Applications in Biological Science and Medicine; Setubal, et al. (eds 1997) Introduction to Computational Molecular Biology; Misener and Krawetz (eds, 2000) Bioinformatics: Methods and Protocols; Higgins and Taylor (eds., 2000) Bioinformatics: Sequence, Structure, and Databanks: A Practical



Approach; Brown (2001) Bioinformatics: A Biologist's Guide to Biocomputing and the Internet; Han and Kamber (2000) Data Mining: Concepts and Techniques (2000); and Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes.

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for lung cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment, the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may



be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results.

Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC,



MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

### **Characteristics of lung cancer-associated proteins**

Lung cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the lung cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or dysregulated cellular processes (see, e.g., Alberts (ed. 1994) Molecular Biology of the Cell (3d ed.). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more structural motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2



domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of amino acid sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (see, e.g., Bateman, et al (2000) *Nuc. Acids Res.* 28:263-266; Sonnhammer, et al. (1997) *Proteins* 28:405-420; Bateman, et al. (1999) *Nuc. Acids Res.* 27:260-262; and Sonnhammer, et al. (1998) *Nuc. Acids Res.* 26:320-322).

In another embodiment, the lung cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels, pumps, and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 17 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the



localization and number of transmembrane domains within the protein may be predicted (see, e.g., PSORT web site <http://psort.nibb.ac.jp/>).

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF, and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, hormones, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they may mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains may also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Lung cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for extracellular immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ* or in histological analysis. Alternatively, antibodies can also label intracellular proteins, in which case analytical samples are typically permeablized to provide access to intracellular proteins. In addition, some membrane proteins can be processed to release a soluble protein, or to expose a residual fragment. Released soluble proteins may be useful diagnostic markers, processed residual protein fragments may be useful lung markers of disease.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the lung cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins may have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; e.g., if circulating, they often serve to transmit



signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor), an endocrine manner (acting on cells at a distance, e.g., secretion into the blood stream), or exocrine (secretion, e.g., through a duct or to adjacent epithelial surface as sweat glands, sebaceous glands, pancreatic ducts, lacrimal glands, mammary glands, wax producing glands of the ear, etc.). Thus secreted molecules often find use in modulating or altering numerous aspects of physiology. Lung cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood, plasma, serum, or stool tests. Those which are enzymes may be antibody or small molecule targets. Others may be useful as vaccine targets, e.g., via CTL mechanisms.

### Use of lung cancer nucleic acids

As described above, lung cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

The lung cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1A-16, can be fragments of larger genes, i.e., they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the lung cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, et al., *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once a lung cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire lung cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant lung cancer nucleic acid can be further-used as a probe to identify and isolate



other lung cancer nucleic acids, e.g., extended coding regions. It can also be used as a “precursor” nucleic acid to make modified or variant lung cancer nucleic acids and proteins.

The lung cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the lung cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, RNAi, vaccine, and/or antisense applications. Alternatively, the lung cancer nucleic acids that include coding regions of lung cancer proteins can be put into expression vectors for the expression of lung cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to lung cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the lung cancer nucleic acids, i.e., the target sequence (either the target sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by “substantially complementary” herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under appropriate reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally complements of ORFs or whole genes are not used. In some embodiments, nucleic acids of lengths up to hundreds of bases can be used.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a



particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By “immobilized” and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By “non-covalent binding” and grammatical equivalents herein is typically meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By “covalent binding” and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to a biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By “substrate” or “solid support” or other grammatical equivalents herein is meant a material that can be modified for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. Often the substrate may contain discrete individual sites appropriate for individual partitioning and identification. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in US application entitled Reusable Low Fluorescent Plastic Biochip, U.S.



Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g., using linkers as are known in the art; e.g., homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized, and then attached to the surface of the solid support. Either the 5' or 3' terminus may be attached to the solid support, or attachment may be via linkage to an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized *in situ*, using known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affymetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of lung cancer-associated sequences. These assays are typically performed in conjunction with



reverse transcription. In such assays, a lung cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of lung cancer-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, e.g., in Innis, et al. (1990) PCR Protocols, A Guide to Methods and Applications.

In some embodiments, a TaqMan based assay is used to measure expression.

TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (see, e.g., literature provided by Perkin-Elmer, e.g., [www2.perkin-elmer.com](http://www2.perkin-elmer.com)).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4:560, Landegren, et al. (1988) Science 241:1077, and Barringer, et al. (1990) Gene 89:117), transcription amplification (Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173), self-sustained sequence replication (Guatelli, et al. (1990) Proc. Nat. Acad. Sci. USA 87:1874), dot PCR, and linker adapter PCR, etc.

### **Expression of lung cancer proteins from nucleic acids**

In a preferred embodiment, lung cancer nucleic acids, e.g., encoding lung cancer proteins, are used to make a variety of expression vectors to express lung cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (see, e.g., Ausubel, *supra*, and Fernandez and Hoeffler (eds 1999) Gene Expression Systems) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the lung cancer protein. The term "control sequences" refers to DNA



sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

5 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding  
10 linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in  
15 accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the lung cancer protein. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are  
20 not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences may be either constitutive or inducible promoters. The promoters  
25 may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two  
30 organisms, e.g., in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector often contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating



vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez and Hoeffler, *supra*).

5 In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

10 The lung cancer proteins of the present invention are usually produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a lung cancer protein, under the appropriate conditions to induce or cause expression of the lung cancer protein. Conditions appropriate for lung cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

20 Appropriate host cells include yeast, bacteria, archaebacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

25 In a preferred embodiment, the lung cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and adenoviral systems. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (see, e.g., Fernandez and Hoeffler, *supra*). Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.



The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, lung cancer proteins are expressed in bacterial systems. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the lung cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez and Hoeffler, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, lung cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, lung cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*,



*Kluyveromyces fragilis* and *K. lactis*, *Pichia guillermondii*, and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

The lung cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the lung cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the lung cancer protein may be made as a fusion protein to increase expression for affinity purification purposes, or for other reasons. For example, when the lung cancer protein is a lung cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

In a preferred embodiment, the lung cancer protein is purified or isolated after expression. Lung cancer proteins may be isolated or purified in a variety of appropriate ways. Standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the lung cancer protein may be purified using a standard anti-lung cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes (1982) Protein Purification. The degree of purification necessary will vary depending on the use of the lung cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the lung cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, therapeutic entities, for production of antibodies, as transcription or translation inhibitors, etc.

## **Variants of lung cancer proteins**

In one embodiment, the lung cancer proteins are derivative or variant lung cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative lung cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at a particular residue within the lung cancer peptide.

Also included within one embodiment of lung cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three



classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the lung cancer protein, using cassette or PCR mutagenesis or other techniques, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above.

5 However, variant lung cancer protein fragments having up to about 100-150 residues may be prepared by *in vitro* synthesis. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the lung cancer protein amino acid sequence. The variants typically exhibit a similar qualitative biological activity as the naturally occurring analogue, 10 although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is often predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be 15 conducted at the target codon or region and the expressed lung cancer variants screened for the optimal combination of desired activity. Techniques exist for making substitution mutations at predetermined sites in DNA having a known sequence, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of mutants is often done using assays of lung cancer protein activities.

20 Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be occasionally tolerated. Deletions generally range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive 25 at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. Larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of a lung cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution chart provided in the definition section.

30 Variants typically exhibit essentially the same qualitative biological activity and will elicit the same immune response as a naturally-occurring analog, although variants also are selected to modify the characteristics of lung cancer proteins as needed. Alternatively, the



variant may be designed or reorganized such that a biological activity of the lung cancer protein is altered. For example, glycosylation sites may be added, altered, or removed.

Covalent modifications of lung cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a lung cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a lung cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking lung cancer polypeptides to a water-insoluble support matrix or surface for use in a method for purifying anti-lung cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimide.

Other modifications include deamidation of glutamyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of serinyl, threonyl or tyrosyl residues, methylation of the  $\gamma$ -amino groups of lysine, arginine, and histidine side chains (Creighton (1983) Proteins: Structure and Molecular Properties, pp. 79-86), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the lung cancer polypeptide encompassed by this invention is an altered native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended herein to mean adding to or deleting one or more carbohydrate moieties of a native sequence lung cancer polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express lung cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to lung cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native sequence lung cancer polypeptide (for O-linked glycosylation sites). The lung cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the lung cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.



Another means of increasing the number of carbohydrate moieties on the lung cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and in Aplin and Wriston (1981) CRC Crit. Rev. Biochem., pp. 259-306.

5 Removal of carbohydrate moieties present on the lung cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al. (1987) Arch. Biochem. Biophys., 259:52 and by Edge, et al. (1981) Anal. Biochem., 118:131.

10 Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura, et al. (1987) Meth. Enzymol., 138:350.

Another type of covalent modification of lung cancer comprises linking the lung cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent  
15 Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192, or 4,179,337.

Lung cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a lung cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric  
20 molecule comprises a fusion of a lung cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the lung cancer polypeptide. The presence of such epitope-tagged forms of a lung cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the lung  
25 cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a lung cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

30 Various tag polypeptides and their respective antibodies are well known and examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field, et al. (1988) Mol. Cell. Biol. 8:2159-2165); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies



thereto (Evan, et al. (1985) Molecular and Cellular Biology 5:3610-3616); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky, et al. (1990) Protein Engineering 3(6):547-553). Other tag polypeptides include the Flag-peptide (Hopp, et al. (1988) BioTechnology 6:1204-1210); the KT3 epitope peptide (Martin, et al. (1992) Science 255:192-194); tubulin epitope peptide (Skinner, et al. (1991) J. Biol. Chem. 266:15163-15166); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:6393-6397).

Also included are other lung cancer proteins of the lung cancer family, and lung cancer proteins from other organisms, which are cloned and expressed as outlined below.

Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related lung cancer proteins from primates or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include unique areas of the lung cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. PCR reaction conditions are well known in the art (e.g., Innis, PCR Protocols, *supra*).

### **Antibodies to lung cancer proteins**

In a preferred embodiment, when a lung cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the lung cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller lung cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are well known (e.g., Coligan, *supra*; and Harlow and Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of Tables 1A-16 or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal



being immunized. Immunogenic proteins include, e.g., keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Adjuvants include, e.g., Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein (1975) Nature 256:495. In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of the tables, or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding (1986) Monoclonal Antibodies: Principles and Practice, pp. 59-103 ). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovin, or primate origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are typically monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid of the tables or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.



In a preferred embodiment, the antibodies to lung cancer protein are capable of reducing or eliminating a biological function of a lung cancer protein, in a naked form or conjugated to an effector moiety. That is, the addition of anti-lung cancer protein antibodies (either polyclonal or preferably monoclonal) to lung cancer tissue (or cells containing lung cancer) may reduce or eliminate the lung cancer. Generally, at least a 25% decrease in activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the lung cancer proteins are humanized antibodies (e.g., Xenerex Biosciences, Medarex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of a human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. A humanized antibody optimally also will typically comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-329; and Presta (1992) Curr. Op. Struct. Biol. 2:593-596). Humanization can be performed following the method of Winter and co-workers (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-327; Verhoeyen, et al. (1988) Science 239:1534-1536), by substituting rodent CDRs or CDR sequences for corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by corresponding sequence from a non-human species.



Human-like antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom and Winter (1991) J. Mol. Biol. 227:381; Marks, et al. (1991) J. Mol. Biol. 222:581). The techniques of Cole, et al. and Boerner, et al. are also available for the preparation of human monoclonal antibodies (Cole, et al. (1985) Monoclonal Antibodies and Cancer Therapy, p. 77 and Boerner, et al. (1991) J. Immunol. 147(1):86-95). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in nearly all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks, et al. (1992) Bio/Technology 10:779-783; Lonberg, et al. (1994) Nature 368:856-859; Morrison (1994) Nature 368:812-13; Fishwild, et al. (1996) Nature Biotechnology 14:845-51; Neuberger (1996) Nature Biotechnology 14:826; and Lonberg and Huszar (1995) Intern. Rev. Immunol. 13:65-93.

By immunotherapy is meant treatment of lung cancer with an antibody raised against a lung cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. The antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the lung cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, may bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted lung cancer protein.

In another preferred embodiment, the lung cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment may bind the extracellular domain of the lung cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The



antibody may cause down-regulation of the transmembrane lung cancer protein. The antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the lung cancer protein. The antibody may be an antagonist of the lung cancer protein or may prevent activation of a transmembrane lung cancer protein, or may induce or suppress a particular cellular pathway. In some embodiments, when the antibody prevents the binding of other molecules to the lung cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF- $\alpha$ , TNF- $\beta$ , IL-1, INF- $\gamma$ , and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody may belong to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, lung cancer may be treated by administering to a patient antibodies directed against the transmembrane lung cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be various molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of a lung cancer protein. In another aspect the therapeutic moiety may modulate an activity of molecules associated with or in close proximity to a lung cancer protein. The therapeutic moiety may inhibit enzymatic or signaling activity such as protease or collagenase activity associated with lung cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to lung cancer tissue or cells results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with lung cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin, saporin, auristatin, and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against lung cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane lung cancer proteins not only serves to increase the local concentration of therapeutic moiety in the lung cancer



afflicted area, but also serves to reduce deleterious side effects that may be associated with the untargeted therapeutic moiety.

In another preferred embodiment, the lung cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein or other entity which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the lung cancer protein can be targeted within a cell, i.e., the nucleus, an antibody theretomay contain a signal for that target localization, i.e., a nuclear localization signal.

The lung cancer antibodies of the invention specifically bind to lung cancer proteins. By “specifically bind” herein is meant that the antibodies bind to the protein with a  $K_d$  of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most preferably, 0.01  $\mu$ M or better. Selectivity of binding to the specific target and not to related other sequences is also important.

#### **Detection of lung cancer sequence for diagnostic and therapeutic applications**

In one aspect, the RNA expression levels of genes are determined for different cellular states in the lung cancer phenotype. Expression levels of genes in normal tissue (e.g., not undergoing lung cancer), in lung cancer tissue (and in some cases, for varying severities of lung cancer that relate to prognosis, as outlined below), or in non-malignant disease are evaluated to provide expression profiles. A gene expression profile of a particular cell state or point of development is essentially a “fingerprint” of the state of the cell. While two states may have a particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

“Differential expression,” or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g.,



normal versus lung cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart (1996) Nature Biotechnology 14:1675-1680, hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is typically at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the RNA or DNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the lung cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to lung cancer genes, e.g., those identified as being important in a lung cancer or disease phenotype, can be evaluated in a lung cancer diagnostic test. In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes.

The lung cancer nucleic acid probes may be attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In a preferred embodiment nucleic acids encoding the lung cancer protein are detected. Although DNA or RNA encoding the lung cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a lung cancer protein is detected.



Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed *in situ*. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a lung cancer protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, lung cancer proteins, including intracellular, transmembrane, or secreted proteins, find use as markers of lung cancer, e.g., for prognostic or diagnostic purposes. Detection of these proteins in putative lung cancer tissue allows for detection, prognosis, or diagnosis of lung cancer or similar disease, and perhaps for selection of therapeutic strategy. In one embodiment, antibodies are used to detect lung cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the lung cancer protein is detected, e.g., by immunoblotting with antibodies raised against the lung cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the lung cancer protein find use in *in situ* imaging techniques, e.g., in histology (e.g., Asai (ed. 1993) Methods in Cell Biology:



Antibodies in Cell Biology, volume 37. In this method cells are contacted with from one to many antibodies to the lung cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable  
5 label, e.g., multicolor fluorescence or confocal imaging. In another method the primary antibody to the lung cancer protein(s) contains a detectable label, e.g., an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of lung cancer proteins. Many other histological  
10 imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing lung cancer from  
15 blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of lung cancer proteins. Antibodies can be used to detect a lung cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune  
20 response against an endogenous lung cancer protein or vaccine.

In a preferred embodiment, *in situ* hybridization of labeled lung cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including lung cancer tissue and/or normal tissue, are made. *In situ* hybridization (see, e.g., Ausubel, *supra*) is then performed. When comparing the fingerprints between an individual and a standard, the  
25 skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids,  
30 modified proteins and cells containing lung cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to lung cancer, clinical, pathological, or other information, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. Single or multiple



genes may be useful in various combinations. As above, lung cancer probes may be attached to biochips for the detection and quantification of lung cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

5

### Assays for therapeutic compounds

In a preferred embodiment, the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in drug screening  
10 assays or by evaluating the effect of drug candidates on a “gene expression profile” or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, et al. (1998) Science 279:84-8; Heid (1996) Genome Res. 6:986-94.

15

In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified lung cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the lung cancer phenotype or an identified physiological function of a lung cancer protein. As above, this can be done on an individual gene level or  
20 by evaluating the effect of drug candidates on a “gene expression profile”. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

25

Having identified differentially expressed genes herein, a variety of assays may be performed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene with altered regulation in lung cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the lung cancer protein. “Modulation” thus includes an increase or a decrease in gene expression. The preferred amount of modulation will depend on the original change of the  
30 gene expression in normal versus tissue undergoing lung cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in lung cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in lung



cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the lung cancer protein and standard immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene or protein expression monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

In this embodiment, the lung cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more lung cancer-associated sequences, e.g., a polynucleotide sequence set out in the tables. Generally, in a preferred embodiment, a test compound is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate lung cancer, modulate lung cancer proteins, bind to a lung cancer protein, or interfere with the binding of a lung cancer protein and an antibody, substrate, or other binding partner.

The term “test compound” or “drug candidate” or “modulator” or grammatical equivalents as used herein describes a molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the lung cancer phenotype or the expression of a lung cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles of nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a lung cancer phenotype, e.g., to a normal or non-malignant tissue fingerprint. In another embodiment, a modulator induces a lung cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.



In one aspect, a modulator will neutralize the effect of a lung cancer protein. By “neutralize” is meant that activity of a protein and the consequent effect on the cell is inhibited or blocked.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a lung cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a “lead compound”) with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such “combinatorial chemical libraries” are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop, et al. (1994) J. Med. Chem. 37(9):1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) Pept. Prot. Res. 37:487-493, Houghton, et al. (1991) Nature, 354:84-88), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs, et al. (1993) Proc. Nat. Acad. Sci. USA 90:6909-6913), vinylogous polypeptides (Hagihara, et al. (1992) J. Amer. Chem. Soc. 114:6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, et



al. (1992) J. Amer. Chem. Soc. 114:9217-9218), analogous organic syntheses of small compound libraries (Chen, et al. (1994) J. Amer. Chem. Soc. 116:2661), oligocarbamates (Cho, et al. (1993) Science 261:1303), and/or peptidyl phosphonates (Campbell, et al. (1994) J. Org. Chem. 59:658). See, generally, Gordon, et al. (1994) J. Med. Chem. 37:1385, nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn, et al. (1996) Nature Biotechnology 14(3):309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang, et al. (1996) Science 274:1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. The above devices, with appropriate modification, are suitable for use with the present invention. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect modulation of lung cancer gene transcription, polypeptide expression, and polypeptide activity.

High throughput assays for evaluating the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins,



U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g.,  
5 Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate procedures, including sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a  
10 high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or  
15 fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly  
20 useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring  
25 proteins, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that the nucleic acid or peptide consists of essentially random sequences of nucleotides and amino acids, respectively. Since these random peptides (or nucleic acids, discussed below) are often chemically synthesized, they may incorporate a nucleotide or amino acid at any position. The synthetic process can be designed to generate  
30 randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.



In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. In a preferred embodiment, the nucleotides or amino acid residues are  
5 randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc.

Modulators of lung cancer can also be nucleic acids, as defined above.

10 As described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or “biased” random nucleic acids. Digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

15 After a candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence is analyzed. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro* transcription  
20 with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence’s specific binding to a probe. The label also can be an enzyme, such as,  
25 alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin  
30 is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

Nucleic acid assays can be direct hybridization assays or can comprise “sandwich assays”, which include the use of multiple probes, as is generally outlined in U.S. Patent Nos.



5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the lung cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state,



screens are performed to identify agents that bind and/or modulate the biological activity of the gene product, or evaluate genetic polymorphisms.

Genes can be screened for those that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a lung cancer expression pattern leading to a normal expression pattern, or to modulate a single lung cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated lung cancer tissue reveals genes that are not expressed in normal tissue or lung cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for lung cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated lung cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of lung cancer cells, that have an associated lung cancer expression profile. By “administration” or “contacting” herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once a test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., lung cancer or non-malignant tissue may be screened for agents that modulate, e.g., induce or suppress a lung cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on lung cancer activity. By defining such a signature for the lung cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.



Measure of lung cancer polypeptide activity, or of lung cancer or the lung cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the metastatic polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of lung cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian lung cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed *in vitro*. For example, a lung cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the lung cancer polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is typically measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the lung cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is typically detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a lung cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or  $\beta$ -gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed



genes are sometimes referred to herein as “lung cancer proteins.” The lung cancer protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another  
5 embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated  
10 gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the lung cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a lung cancer  
15 protein and a candidate compound, and determining the binding of the compound to the lung cancer protein. Preferred embodiments utilize the human lung cancer protein, although other mammalian proteins may also be used, e.g., for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative lung cancer proteins may be used.

20 Generally, in a preferred embodiment of the methods herein, the lung cancer protein or the candidate agent is non-diffusably bound to an insoluble support, preferably having isolated sample receiving areas (e.g., a microtiter plate, an array, etc.). The insoluble supports may be made of a composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of  
25 screening. The surface of such supports may be solid or porous and of a convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents  
30 and samples. The particular manner of binding of the composition is typically not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition, and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation



sequence when the protein is bound to the support), direct binding to “sticky” or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the lung cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the lung cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the lung cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the lung cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g.,  $^{125}\text{I}$  for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor may be a binding moiety known to bind to the target molecule (i.e., a lung cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically



between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by a test compound. Displacement of the competitor is an indication that the test compound is binding to the lung cancer protein and thus is capable of binding to, and potentially modulating, the activity of the lung cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the lung cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the lung cancer protein.

In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the lung cancer proteins. In one embodiment, the methods comprise combining a lung cancer protein and a competitor in a first sample. A second sample comprises a test compound, a lung cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the lung cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the lung cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native lung cancer protein, but cannot bind to modified lung cancer proteins. The structure of the lung cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a lung cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results.



Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a lung cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising lung cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a lung cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g., hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (e.g., cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate lung cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the lung cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting lung cancer cell division is provided. The method comprises administration of a lung cancer inhibitor. In another embodiment, a method of inhibiting lung cancer is provided. The method may comprise administration of a lung cancer inhibitor. In a further embodiment, methods of treating cells or individuals with lung cancer are provided, e.g., comprising administration of a lung cancer inhibitor.

In one embodiment, a lung cancer inhibitor is an antibody as discussed above. In another embodiment, the lung cancer inhibitor is an antisense molecule.



A variety of cell growth, proliferation, viability, and metastasis assays are known to those of skill in the art, as described below.

*Soft agar growth or colony formation in suspension*

5           Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and  
10          grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of lung cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

15           Techniques for soft agar growth or colony formation in suspension assays are described in Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3<sup>rd</sup> ed.), herein incorporated by reference. See also, the methods section of Garkavtsev, et al. (1996), *supra*, herein incorporated by reference.

20          *Contact inhibition and density limitation of growth*

          Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a  
25          higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (<sup>3</sup>H)-thymidine at saturation density can be used to measure density limitation of growth. See Freshney (1994), *supra*. The transformed cells, when transfected with tumor suppressor genes, regenerate a  
30          normal phenotype and become contact inhibited and would grow to a lower density.

          In this assay, labeling index with (<sup>3</sup>H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a lung cancer-associated sequence and are grown for 24 hours at saturation density in



non-limiting medium conditions. The percentage of cells labeling with (<sup>3</sup>H)-thymidine is determined autoradiographically. See, Freshney (1994), *supra*.

#### *Growth factor or serum dependence*

5 Transformed cells typically have a lower serum dependence than their normal counterparts (see, e.g., Temin (1966) J. Natl. Cancer Inst. 37:167-175; Eagle, et al. (1970) J. Exp. Med. 131:836-879); Freshney, *supra*. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

#### *Tumor specific markers levels*

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g.,  
15 Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman (1992) "Angiogenesis and Cancer" in Sem Cancer Biol.

Various techniques which measure the release of these factors are described in  
20 Freshney (1994), *supra*. Also, see, Unkeless, et al. (1974) J. Biol. Chem. 249:4295-4305; Strickland and Beers (1976) J. Biol. Chem. 251:5694-5702; Whur, et al. (1980) Br. J. Cancer 42:305-312; Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184; Freshney Anticancer Res. 5:111-130 (1985).

#### *Invasiveness into Matrigel*

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate lung cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and  
30 invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.



Techniques described in Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with  $^{125}\text{I}$  and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), *supra*.

#### *Tumor growth in vivo*

Effects of lung cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the lung cancer gene is disrupted or in which a lung cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous lung cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous lung cancer gene with a mutated version of the lung cancer gene, or by mutating the endogenous lung cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi, et al. (1989) Science 244:1288). Chimeric targeted mice can be derived according to Hogan, et al. (1988) Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory and Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Washington, D.C.

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic “nude” mouse (see, e.g., Giovanella, et al. (1974) J. Natl. Cancer Inst. 52:921), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley, et al. (1978) Br. J. Cancer 38:263; Selby, et al. (1980) Br. J. Cancer 41:52) can be used as a host. Transplantable tumor cells (typically about  $10^6$  cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a lung cancer-associated sequences are injected subcutaneously. After a suitable length of time,



preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

## 5 Polynucleotide modulators of lung cancer

### *Antisense and RNAi Polynucleotides*

In certain embodiments, the activity of a lung cancer-associated protein is downregulated, or entirely inhibited, by the use of antisense or an inhibitory polynucleotide, i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a  
10 coding mRNA nucleic acid sequence, e.g., a lung cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their  
15 close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the lung cancer protein mRNA. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

20 Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides.  
25 Sense oligonucleotides can, e.g., be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for lung cancer molecules. A preferred antisense molecule is for a lung cancer sequence in the tables, or for a ligand or activator thereof. Antisense or sense  
30 oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein



is described in, e.g., Stein and Cohen (1988) Cancer Res. 48:2659 and van der Krol, et al. (1988) BioTechniques 6:958).

RNA interference is a mechanism to suppress gene expression in a sequence specific manner. See, e.g., Brumelkamp, et al. (2002) Scienceexpress (21March2002); Sharp (1999) Genes Dev. 13:139-141; and Cathew (2001) Curr. Op. Cell Biol. 13:244-248. In mammalian cells, short, e.g., 21 nt, double stranded small interfering RNAs (siRNA) have been shown to be effective at inducing an RNAi response. See, e.g., Elbashir, et al. (2001) Nature 411:494-498. The mechanism may be used to downregulate expression levels of identified genes, e.g., treatment of or validation of relevance to disease.

### *Ribozymes*

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of lung cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto, et al. (1994) Adv. in Pharmacology 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel, et al. (1990) Nucl. Acids Res. 18:299-304; European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Yamada, et al. (1994) Human Gene Therapy 1:39-45; Leavitt, et al. (1995) Proc. Natl. Acad. Sci. USA 92:699-703; Leavitt, et al. (19994) Human Gene Therapy 5:1151-120; and Yamada, et al. (1994) Virology 205: 121-126).

Polynucleotide modulators of lung cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of lung cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by



formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Thus, in one embodiment, methods of modulating lung cancer in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-lung cancer antibody that reduces or eliminates the biological activity of an endogenous lung cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a lung cancer protein. This may be accomplished in any number of ways. In a preferred embodiment, e.g., when the lung cancer sequence is down-regulated in lung cancer, such state may be reversed by increasing the amount of lung cancer gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous lung cancer gene or administering a gene encoding the lung cancer sequence, using known gene-therapy techniques. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), e.g., as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, e.g., when the lung cancer sequence is up-regulated in lung cancer, the activity of the endogenous lung cancer gene is decreased, e.g., by the administration of a lung cancer antisense or RNAi nucleic acid.

In one embodiment, the lung cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to lung cancer proteins. Similarly, the lung cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify lung cancer antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred embodiment, the antibodies are generated to epitopes unique to a lung cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. The lung cancer antibodies may be coupled to standard affinity chromatography columns and used to purify lung cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the lung cancer protein.

### **Methods of identifying variant lung cancer-associated sequences**

Without being bound by theory, expression of various lung cancer sequences is correlated with lung cancer. Accordingly, disorders based on mutant or variant lung cancer genes may be determined. In one embodiment, the invention provides methods for



identifying cells containing variant lung cancer genes, e.g., determining all or part of the sequence of at least one endogenous lung cancer genes in a cell. In a preferred embodiment, the invention provides methods of identifying the lung cancer genotype of an individual, e.g., determining all or part of the sequence of at least one lung cancer gene of the individual.

5 This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced lung cancer gene to a known lung cancer gene, i.e., a wild-type gene.

10 The sequence of all or part of the lung cancer gene can then be compared to the sequence of a known lung cancer gene to determine if any differences exist. This can be done using known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the lung cancer gene of the patient and the known lung cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

15 In a preferred embodiment, the lung cancer genes are used as probes to determine the number of copies of the lung cancer gene in the genome.

In another preferred embodiment, the lung cancer genes are used as probes to determine the chromosomal localization of the lung cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when  
20 chromosomal abnormalities such as translocations, and the like are identified in the lung cancer gene locus.

### **Administration of pharmaceutical and vaccine compositions**

In one embodiment, a therapeutically effective dose of a lung cancer protein or  
25 modulator thereof, is administered to a patient. By “therapeutically effective dose” herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel, et al. (1992) Pharmaceutical Dosage Forms and Drug Delivery; Lieberman, Pharmaceutical Dosage Forms (vols. 1-3), Dekker, ISBN 0824770846,  
30 082476918X, 0824712692, 0824716981; Lloyd (1999) The Art, Science and Technology of Pharmaceutical Compounding; and Pickar (1999) Dosage Calculations). Adjustments for lung cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration,



drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art.

A “patient” for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the lung cancer proteins and modulators thereof of the present invention can be done in a variety of ways, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the lung cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a lung cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. “Pharmaceutically acceptable acid addition salt” refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. “Pharmaceutically acceptable base addition salts” include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose,



lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms  
5 suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that lung cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, etc.) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by  
10 packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a lung cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions  
15 are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate  
20 and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., Remington's Pharmaceutical Science (15th ed., 1980) and Hardman, et al. (eds. 1996) Goodman and Gilman: The Pharmacological Basis of Therapeutics).

Thus, a typical pharmaceutical composition for intravenous administration would be  
25 about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally  
30 administrable compositions will be known or apparent to those skilled in the art, e.g., Remington's Pharmaceutical Science and Goodman and Gilman, The Pharmacological Basis of Therapeutics, *supra*.



The compositions containing modulators of lung cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a “therapeutically effective dose.” Amounts effective for this use will depend upon the severity of the disease and the general state of the patient’s health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a “prophylactically effective dose.” The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer based, at least in part, upon gene expression profiles. Vaccine strategies may be used, in either a DNA vaccine form, or protein vaccine.

It will be appreciated that the present lung cancer protein-modulating compounds can be administered alone or in combination with additional lung cancer modulating compounds or with other therapeutic agent, e.g., other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, e.g., polynucleotides comprising nucleic acid sequences set forth in the tables, such as antisense or RNAi polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of lung cancer-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo*, or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see, e.g.,



Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 (Berger), Ausubel, et al. (eds. 1999) Current Protocols (supplemented through 1999), and Sambrook, et al. (1989) Molecular Cloning - A Laboratory Manual (2nd ed., Vol. 1-3).

5 In a preferred embodiment, lung cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, lung cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the lung cancer coding regions) can be administered in a gene therapy application. These lung cancer genes can include antisense or inhibitory applications, e.g., as inhibitory RNA or gene  
10 therapy (e.g., for incorporation into the genome) or as antisense compositions.

Lung cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL, and antibody responses.. Such vaccine compositions can include, e.g., lipidated peptides (see, e.g., Vitiello, et al. (1995) J. Clin. Invest. 95:341), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres  
15 (see, e.g., Eldridge, et al. (1991) Molec. Immunol. 28:287-294; Alonso, et al. (1994) Vaccine 12:299-306; Jones, et al. (1995) Vaccine 13:675-681), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi, et al. (1990) Nature 344:873-875; Hu, et al. (1998) Clin Exp Immunol. 113:235-243), multiple antigen peptide systems (MAPs) (see, e.g., Tam (1988) Proc. Natl. Acad. Sci. U.S.A. 85:5409-5413; Tam  
20 (1996) J. Immunol. Methods 196:17-32), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, et al., p. 379 In: Kaufmann (ed. 1996) Concepts in vaccine development; Chakrabarti, et al. (1986) Nature 320:535; Hu, et al. (1986) Nature 320:537; Kieny, et al. (1986) AIDS Bio/Technology 4:790; Top, et al. (1971) J. Infect. Dis. 124:148; Chanda, et al.  
25 (1990) Virology 175:535), particles of viral or synthetic origin (see, e.g., Kofler, et al. (1996) J. Immunol. Methods 192:25; Eldridge, et al. (1993) Sem. Hematol. 30:16; Falo, et al. (1995) Nature Med. 7:649), adjuvants (Warren, et al. (1986) Annu. Rev. Immunol. 4:369; Gupta, et al. (1993) Vaccine 11:293), liposomes (Reddy, et al. (1992) J. Immunol. 148:1585; Rock (1996) Immunol. Today 17:131), or, naked or particle absorbed cDNA (Ulmer, et al. (1993)  
30 Science 259:1745; Robinson, et al. (1993) Vaccine 11:957; Shiver, et al., p. 423 In: Kaufmann (ed. 1996) Concepts in vaccine development; Cease and Berzofsky (1994) Annu. Rev. Immunol. 12:923 and Eldridge, et al. (1993) Sem. Hematol. 30:16). Toxin-targeted



delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff, et. al. (1990) Science 247:1465 as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode lung cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, et al. (1991) Nature 351:456-460. A wide variety of other vectors useful for therapeutic administration or immunization e.g., adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the



like, will be apparent to those skilled in the art from the description herein (see, e.g., Shata, et al. (2000) Mol Med Today 6:66-71; Shedlock, et al. (2000) J. Leukoc. Biol. 68:793-806; Hipp, et al. (2000) In Vivo 14:571-85).

5 Methods for the use of genes as DNA vaccines are well known, and include placing a lung cancer gene or portion of a lung cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a lung cancer patient. The lung cancer gene used for DNA vaccines can encode full-length lung cancer proteins, but more preferably encodes portions of the lung cancer proteins including peptides derived from the lung cancer protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a  
10 plurality of nucleotide sequences derived from a lung cancer gene. For example, lung cancer-associated genes or sequence encoding subfragments of a lung cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular  
15 epitopes.

In a preferred embodiment, DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the lung cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

20 In another preferred embodiment lung cancer genes find use in generating animal models of lung cancer. When the lung cancer gene identified is repressed or diminished in metastatic tissue, gene therapy technology, e.g., wherein antisense or inhibitory RNA directed to the lung cancer gene will also diminish or repress expression of the gene. Animal models of lung cancer find use in screening for modulators of a lung cancer-associated sequence or  
25 modulators of lung cancer. Similarly, transgenic animal technology including gene knockout technology, e.g., as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the lung cancer protein. When desired, tissue-specific expression or knockout of the lung cancer protein may be necessary.

It is also possible that the lung cancer protein is overexpressed in lung cancer. As  
30 such, transgenic animals can be generated that overexpress the lung cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene.



Animals generated by such methods will find use as animal models of lung cancer and are additionally useful in screening for modulators to treat lung cancer.

### **Kits for Use in Diagnostic and/or Prognostic Applications**

5           For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In diagnostic and research applications such kits may include at least one of the following: assay reagents, buffers, lung cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, RNAi, dominant negative lung cancer polypeptides or polynucleotides, small molecule inhibitors of  
10 lung cancer-associated sequences, etc. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

          In addition, the kits may include instructional materials containing instructions (e.g., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. A medium  
15 capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

          The present invention also provides for kits for screening for modulators of lung  
20 cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a lung cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing lung cancer-associated activity. Optionally, the kit contains biologically active lung cancer protein. A wide variety of kits and components can be prepared according to the present  
25 invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes typically will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.



## EXAMPLES

## Example 1: Gene Chip Analysis

5 Molecular profiles of various normal and cancerous tissues were determined and analyzed using gene chips. RNA was isolated and gene chip analysis was performed as described (Glynne, et al. (2000) Nature 403:672-676; Zhao, et al. (2000) Genes Dev. 14:981-993).



Tables 1A and 1B were previously filed on April 18, 2001 in USSN 60/284,770 (18501-001500US) and on November 29, 2001 in USSN 60/334,370 (18501-001520US)

5	Table 1A				
	Pkey	ExAccn	UnigenelD	Unigene Title	70% chron/90% NL 70% SQAD/90% NL
	100134	D13264	Hs.49	macrophage scavenger receptor 1	1.61 0.74
	100780	HG3731-HT4001		""Immunoglobulin Heavy Chain, Vdjrc Reg	2.68 3.28
10	100971	J02874	Hs.83213	fatty acid binding protein 4; adipocyte	1.96 0.14
	101088	L05568	Hs.553	solute carrier family 6 (neurotransmitte	0.79 0.07
	101102	L07594	Hs.79059	transforming growth factor; beta recepto	2.55 1
	101168	L15388	Hs.211569	G protein-coupled receptor kinase 5	0.88 0.27
	101277	L38486	Hs.118223	microfibrillar-associated protein 4	0.89 0.26
15	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.59 0.29
	101336	L49169	Hs.75678	FBJ murine osteosarcoma viral oncogene h	1.15 0.41
	101345	L76380	Hs.152175	calcitonin receptor-like	0.81 0.31
	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a l	1.31 0.77
	101764	M80563	Hs.81256	S100 calcium-binding protein A4 (calcium	1.44 0.82
20	101771	M81750	Hs.153837	myeloid cell nuclear differentiation ant	0.96 0.45
	101842	M93221	Hs.75182	mannose receptor; C type 1	1.27 0.37
	102283	U31384	Hs.83381	guanine nucleotide binding protein 11	1.04 0.3
	102363	U39447	Hs.198241	amine oxidase; copper containing 3 (vasc	0.96 0.26
	102507	U52154	Hs.193044	potassium inwardly-rectifying channel; s	2.81 3.45
25	102698	U75272	Hs.1867	progastricsin (pepsinogen C)	0.95 0.23
	103025	X54131	Hs.123641	protein tyrosine phosphatase; receptor t	1.62 0.21
	103280	X79981	Hs.76206	cadherin 5; VE-cadherin (vascular epithe	0.9 0.41
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2	1.27 0.49
	103541	Z11697	Hs.79197	CD83 antigen (activated B lymphocytes; i	1.86 1
30	103554	Z18951	Hs.74034	caveolin 1; caveolae protein; 22kD	1.27 0.47
	104212	AB002298	Hs.173035	KIAA0300 protein	1.17 0.16
	104691	AA011176	Hs.37744	ESTs	1.08 0.35
	104825	AA035613	Hs.141883	ESTs	0.75 0.27
	104857	AA043219	Hs.19058	ESTs	2.6 3.3
35	104865	AA045136	Hs.22575	ESTs	1.23 0.49
	104989	AA102098	Hs.118615	ESTs	0.63 0.32
	105729	AA292694	Hs.3807	ESTs; Weakly similar to PHOSPHOLEMMAN PR	0.86 0.34
	105847	AA398606	Hs.32241	ESTs	1.32 0.4
	105894	AA400979	Hs.25691	calcitonin receptor-like receptor activi	0.78 0.28
40	106490	AA451861	Hs.115537	ESTs; Weakly similar to dipeptidase prec	1.2 0.47
	106536	AA453997	Hs.23804	ESTs	0.82 0.15
	106605	AA457718	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	0.99 0.07
	106667	AA461086	Hs.16578	ESTs	1.17 0.4
	106773	AA478109	Hs.188833	ESTs	1.46 0.43
45	106797	AA478962	Hs.169943	ESTs	1.18 0.32
	106844	AA485055	Hs.158213	sperm associated antigen 6	0.98 0.51
	106870	AA487576	Hs.26530	serum deprivation response (phosphatidyl	1.05 0.14
	106954	AA496980	Hs.204038	ESTs	1.25 0.33
	107054	AA600150	Hs.14366	ESTs	1.11 0.4
50	107292	T30407	Hs.4789	ESTs; Weakly similar to oxidative-stress	1.07 2.58
	107994	AA036811	Hs.165030	ESTs	0.7 0.21
	107997	AA037388	Hs.82223	Human DNA sequence from clone 141H5 on c	1.02 0.48
	108041	AA041552	Hs.61957	ESTs	1.44 0.51
	108087	AA045709	Hs.40545	ESTs	1.98 1
55	108382	AA074885	Hs.67726	macrophage receptor with collagenous str	1.52 0.72
	108435	AA078787	Hs.194101	ESTs	2.53 1.53
	108480	AA081093	Hs.68055	ESTs	1.56 0.48
	109252	AA194830	Hs.85944	ESTs	2.69 3.18
	109550	F01534	Hs.26981	ESTs	1.19 0.65
60	109613	F03031	Hs.27519	ESTs	1.01 0.29
	109837	H00656	Hs.29792	ESTs	0.81 0.15
	109893	H04768	Hs.30484	ESTs	1.44 0.32
	109984	H09594	Hs.10299	ESTs	0.62 0.14
	110099	H16568	Hs.23748	ESTs	1.01 0.28
65	110837	N30796	Hs.17424	ESTs; Weakly similar to semaphorin F [H.	1.1 0.22
	111247	N69825	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	1.26 0.26
	111341	N80935	Hs.22483	ESTs	1.57 0.52
	111510	R07856	Hs.16355	ESTs	3.96 1
	111737	R25410	Hs.9218	ESTs	0.97 0.24
70	113195	T57112		""yc20g11.s1 Stratagene lung (#937210)	1.22 0.35
	113238	T62979	Hs.189813	ESTs	2.27 0.45
	113540	T90496	Hs.16757	ESTs	1.06 0.22
	113552	T90889	Hs.16026	ESTs	1.16 0.42
	113606	T93093	Hs.17125	ESTs	1.48 0.7
75	113695	T96965	Hs.17948	ESTs	1.54 0.28
	113946	W84753	Hs.37896	ESTs	1.79 0.72
	114251	Z39898	Hs.21948	ESTs	1.95 0.25
	114359	Z41589	Hs.153483	ESTs; Moderately similar to H1 chloride	1.42 0.13
	115230	AA278300	Hs.182980	ESTs	2.62 0.42
80	115279	AA279760	Hs.63671	ESTs	1.79 0.91
	115566	AA398083	Hs.43977	ESTs	0.86 0.2
	115965	AA446661	Hs.173233	ESTs	0.79 0.04
	116166	AA461556	Hs.202949	KIAA1102 protein	2.29 0.68
	116279	AA486073	Hs.57362	ESTs	2.27 0.78
	117023	H88157	Hs.41105	ESTs	1.36 0.16



5	117209	H99959	Hs.42768	ESTs	1.46	0.48
	118901	N90719	Hs.94445	ESTs	1.51	1
	118981	N93839	Hs.39288	ESTs	1.34	0.48
	119073	R32894	Hs.45514	v-els avian erythroblastosis virus E26 o	1.14	0.27
	119221	R98105		""yr30g11.s1 Soares fetal liver spleen	1.32	0.53
10	119824	W74536	Hs.184	advanced glycosylation end product-speci	1	0.19
	119861	W80715		ESTs; Moderately similar to !!!! ALU SUB	1.83	0.45
	120041	W92775	Hs.59368	ESTs	1.23	0.55
	120132	Z38839	Hs.125019	ESTs; Highly similar to KIAA0886 protein	0.91	0.37
	120467	AA251579	Hs.187628	ESTs	1.87	1.91
15	121314	AA402799	Hs.182538	ESTs	1.3	0.31
	121643	AA417078	Hs.193767	ESTs	2.31	0.68
	121690	AA418074	Hs.110286	ESTs	1.47	0.51
	122633	AA454080	Hs.34853	inhibitor of DNA binding 4; dominant neg	1.31	0.63
	123978	C20653	Hs.170278	ESTs	1.52	0.32
20	124214	H58608	Hs.151323	ESTs	0.93	0.35
	124357	N22401		""yw37g07.s1 Morton Fetal Cochlea Homo	1.29	1
	124438	N40188	Hs.102550	ESTs	1.36	0.7
	125167	W45560	Hs.102541	ESTs	1.46	0.69
	125174	W51835	Hs.231082	EST	3.07	3.76
25	125422	AA903229	Hs.153717	ESTs	1.34	0.3
	125561	AI417667	Hs.22978	ESTs	1.89	0.63
	125831	D60988		""HUM145B09B Clontech human fetal brain	0.94	0.36
	127002	R35380	Hs.24979	ESTs	3.02	4.06
	127307	AA369367	Hs.126712	ESTs; Weakly similar to pL2 hypothetica	1.01	0.69
30	127609	AA622559	Hs.150318	ESTs	1.21	0.32
	127959	AI302471	Hs.124292	ESTs	2.5	1
	128458	D52193	Hs.56340	ESTs	1.13	0.33
	128624	AA479209	Hs.102647	ESTs	1.45	0.58
	128789	AA486567	Hs.105695	ESTs	1.1	0.34
35	128798	AF014958	Hs.105938	chemokine (C-C motif) receptor-like 2	1.16	0.55
	128952	R51076	Hs.107361	ESTs; Highly similar to Rap2 interacting	2.04	2.4
	129057	X62466	Hs.214742	CDW52 antigen (CAMPATH-1 antigen)	1.77	0.73
	129210	AA401654	Hs.202949	KIAA1102 protein	1.11	0.36
	129240	W24360	Hs.237868	interleukin 7 receptor	0.91	0.41
40	129402	T63781		""yc21g01.s1 Stratagene lung (#937210)	1.36	0.43
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1	0.67	0.08
	129593	AA487015	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	1.3	0.42
	129626	AA447410	Hs.11712	ESTs; Weakly similar to !!!! ALU SUBFAMI	1.28	0.46
	129699	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ubiqu	1.58	1
45	129898	N48595	Hs.13256	ESTs	1.13	0.53
	129958	L20591	Hs.1378	annexin A3	0.81	0.31
	130273	U59914	Hs.153863	MAD (mothers against decapentaplegic; Dr	0.59	0.22
	130655	N92934	Hs.17409	cysteine-rich protein 1 (intestinal)	1.44	0.76
	130657	T94452	Hs.201591	ESTs	0.96	0.42
50	131061	N64328	Hs.22567	ESTs; Moderately similar to HYPOTHETICAL	1.51	0.45
	131066	F09006	Hs.22588	ESTs	0.97	0.37
	131263	R38334	Hs.24950	regulator of G-protein signalling 5	2.34	2.82
	131589	U52100	Hs.29191	epithelial membrane protein 2	1.2	0.62
	131686	AA157428	Hs.30687	Grb2-associated binder 2	0.95	0.38
55	131751	H18335	Hs.31562	ESTs	1.47	0.52
	132430	T23630	Hs.258675	EST	1.86	2.09
	132476	N67192	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	1.73	0.58
	132836	F09557	Hs.57929	slit (Drosophila) homolog 3	0.91	0.29
	133120	X64559	Hs.65424	tetranectin (plasminogen-binding protein	0.82	0.2
60	133488	D45370	Hs.74120	adipose specific 2	1.29	0.48
	133565	H57056	Hs.204831	ESTs	2.25	0.57
	133651	U97105	Hs.173381	dihydropyrimidinase-like 2	1.65	0.62
	133835	AA059489	Hs.76640	ESTs; Highly similar to RGC-32 [R.norveg	1.16	0.34
	133978	W73859	Hs.78061	transcription factor 21	0.79	0.27
65	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec	0.99	0.28
	134299	AA487558	Hs.8135	ESTs	1.02	0.46
	134300	U81984	Hs.166082	endothelial PAS domain protein 1.	0.86	0.42
	134323	AA028976	Hs.8175	Homo sapiens mRNA; cDNA DKFZp564M0763 (f	1.19	0.27
	134343	D50683	Hs.82028	transforming growth factor; beta recepto	1.21	0.67
70	134417	D87969	Hs.82921	solute carrier family 35 (CMP-sialic aci	1.28	1
	134561	U76421	Hs.85302	adenosine deaminase; RNA-specific; B1 (h	2.12	0.55
	134624	W67147	Hs.8700	deleted in liver cancer 1	2.35	2.74
	134696	H88354	Hs.8861	ESTs	1.35	0.33
	134749	L10955	Hs.89485	carbonic anhydrase IV	0.89	0.2
75	134786	L06139	Hs.89640	TEK tyrosine kinase; endothelial (venous	0.48	0.21
	134869	T35288	Hs.90421	ESTs; Moderately similar to !!!! ALU SUB	2.14	2.64
	135346	M21056	Hs.992	phospholipase A2; group IB (pancreas)	0.63	0.13
	100113	D00591	Hs.84746	Chromosome condensation 1	1	2.15
	100147	D13666	Hs.136348	Homo sapiens mRNA for osteoblast specifi	0.5	2
80	100280	D42085	Hs.155314	KIAA0095 gene product	1.02	1.39
	100335	D63391	Hs.6793	platelet-activating factor acetylhydrola	1	5.58
	100360	D78335	Hs.75939	Uridine monophosphate kinase	0.91	2.04
	100372	D79997	Hs.184339	KIAA0175 gene product	0.75	2.03
	100486	HG1112-HT1112		TIGR: ras-like protein TC4	1.09	1.93
85	100559	HG2197-HT2267		"collagen, type VII, alpha 1"	0.97	3.6
	100576	HG2290-HT2386		"calcitonin/alpha-CGRP, alt. transcript	1	1
	100668	HG2981-HT3938		"TIGR: CD44 (epican, alt. transcript 12	0.85	1.9
	100906	HG4716-HT5158		Guanosine 5'-Monophosphate Synthase	1.18	2.29
	100930	HG721-HT4827		"TIGR: placental protein 14, endometrial	1	1.45



5	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	0.84	2.6
	101031	J05070	Hs.151738	"Matrix metalloproteinase 9 (gelatinase	0.77	1.52
	101111	L08424	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
	101124	L10343	Hs.112341	"Protease inhibitor 3, skin-derived (SKA	0.62	2.67
	101175	L18920	Hs.36980	"Melanoma antigen, family A, 2"	1	1
10	101204	L24203	Hs.82237	Ataxia-telangiectasia group D-associated	0.74	4.1
	101431	M19888	Hs.1076	Small proline-rich protein 1B (cornifin)	0.85	2.51
	101448	M21389	Hs.195850	keratin 5 (epidermolysis bullosa simplex	0.61	8.83
	101511	M27826	Hs.267319	Endogenous retroviral protease	1.03	1.13
	101526	M29540	Hs.220529	Carcinoembryonic antigen-related cell ad	1.07	4.61
15	101548	M31328	Hs.71642	"Guanine nucleotide binding protein (G p	0.97	1.13
	101625	M57293		"Human parathyroid hormone-related pepti	1	1
	101649	M60047	Hs.1690	Heparin-binding growth factor binding pr	1	2.7
	101724	M69225	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	1	8.98
	101748	M76482	Hs.1925	Desmoglein 3 (pemphigus vulgaris antigen	1	2.78
20	101759	M80244	Hs.184601	"Solute carrier family 7 (cationic amino	1.07	2.45
	101804	M86699	Hs.169840	TTK protein kinase	1	1
	101806	M86757	Hs.112408	S100 calcium-binding protein A7 (psorias	0.74	1.76
	101809	M86849		"Homo sapiens connexin 26 (GJB2) mRNA, c	1	7
	101845	M93426	Hs.78867	"Protein tyrosine phosphatase, receptor-	1	1
25	101851	M94250	Hs.82045	Midkine (neurite growth-promoting factor	1.13	2.6
	102083	U10323	Hs.75117	"Interleukin enhancer binding factor 2,	1.03	1.61
	102154	U17760	Hs.75517	"Laminin, beta 3 (nicein (125kD), kalini	0.94	3.62
	102193	U20758	Hs.313	secreted phosphoprotein 1 (osteopontin;	0.34	4.59
	102305	U33286	Hs.90073	chromosome segregation 1 (yeast homolog)	1.45	2.97
30	102348	U37519	Hs.87539	Aldehyde dehydrogenase 8	0.52	2.25
	102581	U61145	Hs.77256	Enhancer of zeste (Drosophila) homolog 2	0.91	2.46
	102610	U65011	Hs.30743	Preferentially expressed antigen in mela	1	3.88
	102623	U66083	Hs.37110	"Melanoma antigen, family A, 9 (MAGE-9)"	1	1
	102669	U71207	Hs.29279	Eyes absent (Drosophila) homolog 2	1	1
35	102696	U74612	Hs.239	Forkhead box M1	1.06	2.77
	102829	U91618	Hs.80962	Neurotensin	1	1
	102888	X04741	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.13	2.59
	102913	X07696	Hs.80342	keratin 15	0.7	4.72
	102915	X07820	Hs.2258	Matrix Metalloproteinase 10 (Stromolysin	1.15	3.35
40	102963	X15943	Hs.37058	"Calcitonin/calcitonin-related polypepti	1	1
	103021	X53587	Hs.85266	"Integrin, beta 4"	1.38	2.34
	103036	X54925	Hs.83169	Matrix metalloprotease 1 (interstitial c	1	14.93
	103058	X57348	Hs.184510	Stratifin	1.25	4.17
	103060	X57766	Hs.155324	matrix metalloproteinase 11 (stromelysin	1	1.72
45	103119	X63629	Hs.2877	"Cadherin 3, P-cadherin (placental)"	1.16	7.38
	103206	X72755	Hs.77367	monokine induced by gamma interferon	0.71	1.48
	103242	X76342	Hs.389	"Alcohol dehydrogenase 7 (class IV), mu	1	1
	103312	X82693	Hs.3185	"Lymphocyte antigen 6 complex, locus D;	0.92	1.28
	103478	Y07755	Hs.38991	S100 calcium-binding protein A2	1.05	5.81
50	103558	Z19574	Hs.2785	keratin 17	0.65	6.68
	103576	Z26317	Hs.2631	Desmoglein 2	0.79	1.73
	103587	Z29083	Hs.82128	5T4 Oncofetal antigen	1	3.93
	103594	Z31560	Hs.816	"SRY (sex determining region Y)-box 2, p	0.71	7.23
	103768	AA089997		"ESTs, Highly similar to integral membra	0.99	1.8
55	104158	AA454908	Hs.8127	KIAA0144 gene product	0.96	1.29
	104558	R56678	Hs.88959	Human DNA sequence from clone 967N21 on	1.23	7.23
	104689	AA010665		ESTs	0.96	2.11
	104733	AA019498	Hs.23071	ESTs	1.18	1.88
	104906	AA055809	Hs.26802	Protein kinase domains containing protei	1.11	3.15
60	104978	AA088458	Hs.19322	ESTs; Weakly similar to !!!! ALU SUBFAMI	1.64	2.89
	105012	AA116036	Hs.9329	"Homo sapiens mRNA for fls353, complete	1.19	3.91
	105175	AA186804	Hs.25740	ESTs; Weakly similar to unknown [S.cerev	0.9	4.63
	105263	AA227926	Hs.6682	ESTs	0.95	2.87
	105298	AA233459	Hs.26369	ESTs	1	1.13
65	105312	AA233854	Hs.23348	S-phase kinase-associated protein 2 (p45	1.32	3.01
	105719	AA291644	Hs.36793	Hypothetical protein FLJ23188	1.28	2.31
	105743	AA293300	Hs.9598	ESTs	1	1
	106012	AA411621	Hs.8895	ESTs; same as BFH6?	0.94	2.04
	106231	AA429571	Hs.38002	KIAA1355 protein	1.04	1.5
70	106540	AA454607	Hs.38114	Hypothetical protein FLJ11100	1.26	2.26
	106575	AA456039	Hs.105421	ESTs	1	2
	106632	AA459897	Hs.11950	GPI-anchored metastasis-associated prote	0.87	1.32
	106727	AA465342	Hs.34045	Hypothetical protein FLJ20764	0.87	1.59
	106906	AA490237	Hs.222024	Transcription factor BMAL2 (cycle-like f	0.61	1.6
75	107059	AA608545	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	0.48	2.67
	107104	AA609786	Hs.15243	Nucleolar protein 1 (120kD)	1.01	1.44
	107151	AA621169	Hs.8687	ESTs; procollagen I-N proteinase	0.97	2.89
	107284	S74039	Hs.291904	Accessory proteins BAP31/BAP29	1.15	3.65
	107901	AA026418	Hs.91539	ESTs	0.72	3.44
80	107922	AA028028	Hs.61460	Ig superfamily receptor LNIR precursor	1	2.48
	107932	AA029317	Hs.18878	Hypothetical protein FLJ21620	1	1
	108695	AA121315	Hs.70823	KIAA1077 protein	0.91	3.53
	108857	AA133250	Hs.62180	ESTs	1	1
	108860	AA133334	Hs.129911	ESTs	0.73	7.3
85	108990	AA152296	Hs.72045	ESTs	1	1
	109166	AA179845	Hs.73625	"RAB6 interacting, kinesin-like (rabkine	1	4.55
	109424	AA227919	Hs.85962	Hyaluronan synthase 3	1	1.28
	109665	F05012	Hs.27027	Hypothetical protein DKFZp762H1311	1.42	2
	109970	H09281	Hs.13234	ESTs	1.13	2.16



5	110015	H10998	Hs.7164	A disintegrin and metalloproteinase doma	0.84	1.95
	110156	H18957	Hs.4213	ESTs	0.94	1.41
	110561	H59617	Hs.5199	HSPC150 protein similar to ubiquitin-con	0.91	3.18
	111223	N68921	Hs.34806	ESTs; Weakly similar to neogenin [H.sapi	0.91	3.13
	111345	N89820	Hs.14559	Hypothetical protein FLJ10540	1	1.25
10	111876	R38239	Hs.293246	"ESTs, Weakly similar to putative p150 [	0.83	1.27
	111902	R39191	Hs.109445	KIAA1020 protein	0.91	0.91
	112244	R51309	Hs.70823	KIAA1077 protein	0.77	3.01
	112973	T17271		"cDNA FLJ13308 fis, clone OVARC1001436,	1	1
	112989	T23482	Hs.89981	"Diacylglycerol kinase, zeta (104kD)"	0.55	1.03
15	113047	T25867	Hs.7549	ESTs	0.87	2
	113095	T40920	Hs.126733	ESTs	1	1
	113531	T90345	Hs.16740	Hypothetical protein FLJ11036	0.42	1.44
	113970	W86748	Hs.8109	ESTs	1.17	1.73
	114346	Z41450	Hs.130489	"ATPase, aminophospholipid transporter-I	0.86	0.82
20	114407	AA010188	Hs.103305	ESTs	0.8	1.88
	114471	AA028074	Hs.104613	RP42 homolog	1.06	1.34
	114509	AA043551	Hs.101799	KIAA1350 protein	1.82	2.32
	115060	AA253214	Hs.198249	"Gap junction protein, beta 5 (connexin	0.79	1.49
	115091	AA255900	Hs.184523	KIAA0965 protein	0.72	1.92
25	115123	AA256642	Hs.236894	"ESTs, High sim to LRP1_hu low density l	0.59	1.97
	115291	AA279943	Hs.122579	ESTs	1	1.25
	115506	AA292537	Hs.45207	Hypothetical protein KIAA1335	1.15	1.48
	115522	AA331393	Hs.47378	ESTs	0.5	3.29
	115536	AA347193	Hs.62180	ESTs	1	1
30	115697	AA411502	Hs.63325	Homo sapiens type II membrane serine pro	1	6.53
	115909	AA436666	Hs.59761	ESTs	1	6.98
	115978	AA447522	Hs.69517	Differentially expressed in Fanconi anem	1	2.31
	116028	AA452112	Hs.42644	thioredoxin-like	0.99	1.68
	116107	AA456968	Hs.92030	ESTs	1.14	1.8
35	116134	AA460246	Hs.50441	CGI-04 protein	1.11	1.86
	116157	AA461063	Hs.44298	Hypothetical protein	0.99	1.9
	116158	AA461187	Hs.61762	Hypoxia-inducible protein 2	0.44	0.86
	116335	AA495830	Hs.87013	"Homo sapiens cDNA FLJ10238 fis, clone H	0.62	3.89
	116483	C14092	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.04	2.36
40	117320	N23239	Hs.211092	LUNX protein; PLUNC(palate lung & nasal	0.51	0.64
	117557	N33920	Hs.44532	Diubiquitin	1.11	2.63
	117693	N40939	Hs.112110	PTD007 protein	0.98	1.79
	117881	N50073	Hs.260622	Butyrate-induced transcript 1	1	1.43
	118368	N64339	Hs.48956	ESTs	0.67	2.86
45	118566	N68558	Hs.42824	Hypothetical protein FLJ10718	1.21	0.83
	118695	N71781	Hs.50081	KIAA1199 see CVA7.doc	0.88	1.63
	119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical pro	1	1
	119845	W79920	Hs.58561	G protein-coupled receptor 87	1	1
	120102	W95428	Hs.132927	"ESTs, Moderately similar to p53 regulat	1	1
50	120104	W95477	Hs.180479	ESTs	0.69	3.07
	120486	AA253400	Hs.137569	Tumor protein 63 kDa with strong homolog	1.08	12.05
	120859	AA350158	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
	120880	AA360240	Hs.97019	EST	1	1
	120948	AA397822	Hs.104650	Hypothetical protein FLJ10292	1.04	2.15
55	120983	AA398209	Hs.97587	EST	1	1
	121362	AA405500	Hs.97932	Chondromodulin I precursor	1	1
	121369	AA405657	Hs.128791	CGI-09 protein	1	1.8
	121791	AA423978	Hs.293317	"ESTs, Weakly similar to JM27 [H.sapiens	1	1
	123005	AA479726	Hs.105577	ESTs	1	1
60	123044	AA481549	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	0.95	1.88
	123160	AA488687	Hs.284235	ESTs	1.59	4.98
	123479	AA599469	Hs.135056	clone RP5-850E9 on chromosome 20	1.19	1.64
	123571	AA608956	Hs.112619	"ESTs, Weakly similar to PQ0109 Purkinje	1.03	1.14
	123829	AA620697	Hs.112208	XAGE-1 protein	1.39	2.2
65	124006	D60302	Hs.108977	ESTs	1	4.85
	124059	F13673	Hs.99769	ESTs	1.49	8.62
	124960	T15386	Hs.194766	Seizure related gene 6 (mouse)-like	0.76	0.77
	125218	W73561	Hs.110024	NADH:ubiquinone oxidoreductase MLRQ subu	1.33	1.77
	125453	R06041	Hs.18048	"Melanoma antigen, family A, 10"	0.8	1.42
70	125759	AA425587	Hs.82226	Glycoprotein (transmembrane) nmb	1.52	2.26
	125972	AA434562	Hs.35406	"ESTs, Highly similar to unnamed protein	1.05	2.48
	125994	H55782	Hs.270799	EST	1	1.95
	126395	N70192	Hs.278956	Hypothetical protein FLJ12929	1	1.35
	126645	AI167942	Hs.61635	STEAP1 (Homo sapiens BAC clone RG041D11	1	2.23
75	127221	AI354332	Hs.72365	ESTs	0.73	3.27
	127479	AA513722	Hs.179729	collagen; type X; alpha 1 (Schmid metaph	0.51	1.94
	128192	AI204246		KIAA1085 protein	1.8	3.16
	128610	L38608	Hs.10247	activated leucocyte cell adhesion molecu	0.89	0.97
	128777	U46006	Hs.10526	Cysteine and glycine-rich protein 2	1	1
80	128924	AA234962	Hs.26557	Plakophilin 3	1.3	2.97
	129041	H58873	Hs.169902	"Solute carrier family 2 (facilitated gl	0.84	2.04
	129099	H50398	Hs.108660	"ATP-binding cassette, sub-family C (CFT	0.87	1.04
	129404	AA172056	Hs.111128	ESTs	1	1
	129466	L42583		"Genbank Homo sapiens keratin 6 isoform	0.72	12.67
85	129605	S72493	Hs.115947	Keratin 16 (focal non-epidermolytic palm	0.92	1.5
	129628	U26727	Hs.1174	"Cyclin-dependent kinase inhibitor 2A (m	0.85	1.93
	130023	X13461	Hs.239600	Calmodulin-like 3	0.84	1.22
	130080	X14850	Hs.147097	"H2A histone family, member X"	0.98	1.96
	130385	AA126474	Hs.155223	stanniocalcin 2	1	1



	130410	V01514	Hs.155421	Alpha-fetoprotein	0.63	0.63
	130441	U35835	Hs.301387	"Human DNA-PK mRNA, partial cds"	1.15	3.65
	130482	L32866	Hs.1578	Baculoviral IAP repeat-containing 5 (sur	1	1.88
5	130553	AA430032	Hs.252587	Pituitary tumor-transforming 1	0.92	1.96
	130577	M35410	Hs.162	Insulin-like growth factor binding prote	1.17	4.7
	130627	L23808	Hs.1695	Matrix metalloproteinase 12 (macrophage	0.69	4.05
	130800	AA223386	Hs.19574	ESTs; Weakly similar to katanin p80 subu	1.13	2.41
	130939	AA598689	Hs.21400	ESTs	0.8	0.89
10	131046	X02530	Hs.2248	INTERFERON-GAMMA INDUCED PROTEIN PRECURS	0.8	1.15
	131244	D38076	Hs.24763	RAN binding protein 1	1.13	1.85
	131877	J04088	Hs.156346	Topoisomerase (DNA) II alpha (170kD)	1	1
	131927	AA461549	Hs.34780	"Doublecortx; lissencephaly, X-linked (	0.81	0.62
	131965	W90146	Hs.35962	ESTs	0.74	3.27
15	131978	D80008	Hs.36232	KIAA0186 gene product	1	1
	132354	L05187	Hs.211913	Small proline-rich protein 1A	0.69	1.43
	132543	AA417152	Hs.5101	ESTs; Highly similar to protein regulati	0.79	4.27
	132632	N59764	Hs.5398	guanine-monophosphate synthetase	1	1.08
	132653	U31201	Hs.54451	"laminin gamma2 chain gene (LAMC2), exon	1	1
20	132659	Z75190	Hs.54481	"Low density lipoprotein receptor-relate	0.89	0.89
	132710	W93726	Hs.55279	"Serine (or cysteine) proteinase inhibit	0.64	4.41
	132758	W52432	Hs.56105	"ESTs, Weakly similar to WDNM RAT WDNM1	1.55	2.08
	132767	L05188	Hs.231622	Small proline-rich protein 2B	0.83	1.66
	132816	M74542	Hs.575	Aldehyde dehydrogenase 3	0.55	0.55
25	132990	AA458761	Hs.18387	transcription factor AP-2 alpha (activat	1	3.53
	133070	U69611	Hs.64311	"A disintegrin and metalloproteinase dom	1.16	2
	133282	U52960	Hs.286145	"SRB7 (suppressor of RNA polymerase B, y	1	2.7
	133317	AA215299	Hs.70830	U6 snRNA-associated Sm-like protein LSM7	0.95	1.42
	133370	AA156897	Hs.72157	Homo sapiens mRNA; cDNA DKFZp564I1922	1.12	2.55
30	133391	X57579	Hs.727	H.sapiens activin beta-A subunit (exon 2	1.65	1.76
	133832	H03387	Hs.241305	estrogen-responsive B box protein (EBBP)	1.02	1.39
	134032	Z81326	Hs.78589	"Serine (or cysteine) proteinase inhibit	1	1
	134168	AA398908	Hs.181634	"Homo sapiens cDNA: FLJ23602 fis, clone	0.95	1.53
	134218	AA227480	Hs.80205	Pim-2 oncogene	1.36	2.48
35	134405	R67275	Hs.82772	""collagen, type XI, alpha 1""	0.76	2.86
	134453	X70683	Hs.83484	SRY (sex determining region Y)-box 4	1.89	3.78
	134470	X54942	Hs.83758	CDC28 protein kinase 2	1.82	4.11
	134645	U87459	Hs.167379	"Cancer/testis antigen (NY-ESO-1, CTAG1,	0.82	0.83
	134781	M17183	Hs.89626	Parathyroid hormone-like hormone	1	1
40	135002	U19147	Hs.272484	G antigen 6	1	1
	100040	M97935		AFFX control: STAT1	0.92	1.25
	101201	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin;	2.92	8.5
	101664	M60752	Hs.121017	H2A histone family; member A	1	1
	102025	U03911	Hs.78934	mutS (E. coli) homolog 2 (colon cancer;	0.8	1.61
45	102031	U04898	Hs.2156	RAR-related orphan receptor A	1	1
	102221	U24576		LIM domain only 4	1	1
	102270	U30255	Hs.75888	phosphogluconate dehydrogenase	1.08	1.43
	102339	U37022	Hs.95577	cyclin-dependent kinase 4	0.88	1.32
	102391	U41668	Hs.77494	deoxyguanosine kinase	1.07	1.58
50	103000	X51956	Hs.146580	enolase 2; (gamma; neuronal)	0.91	1.49
	103395	X94754	Hs.119503	methionine-tRNA synthetase	0.89	1.32
	105638	AA281599	Hs.20418	Homo sapiens mRNA for for histone H2B; c	0.91	1.25
	105726	AA292328	Hs.9754	activating transcription factor 5	0.94	1.48
	114841	AA234722	Hs.55408	ESTs; Moderately similar to CALCIUM-DEPE	0.78	1.56
55	115206	AA262491	Hs.186572	ESTs	1	1
	115906	AA436616	Hs.82302	ESTs	0.74	2.52
	119132	R49046	Hs.107911	ATP-binding cassette; sub-family B (MDR/	1.1	1.51
	124163	H30539	Hs.189838	ESTs	1	1
	126487	AA482505	Hs.184601	solute carrier family 7 (cationic amino	1.01	1.46
60	127141	AA307960	Hs.75478	KIAA0956 protein	0.85	1.4
	128034	AA905754	Hs.75103	tyrosine 3-monooxygenase/tryptophan 5-mo	1	1.18
	128609	AA234365	Hs.102456	survival of motor neuron protein interac	1	1.5
	128895	R37753	Hs.106985	ESTs	1.7	2
	130199	Z48579	Hs.172028	a disintegrin and metalloprotease domain	1	1
65	130524	U89995	Hs.159234	forkhead box E1	1	1
	133000	U24152	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	1	1
	133658	M25756	Hs.75426	secretogranin II (chromogranin C)	1	1
	135047	AA460466	Hs.93597	ESTs	1	1
	100053	M27830		AFFX control: 28S ribosomal RNA	0.88	1.53
70	100114	D00596	Hs.82962	thymidylate synthetase	0.68	1.86
	100128	D11094	Hs.61153	proteasome (prosome; macropain) 26S subu	1.29	2.03
	100154	D14657	Hs.81892	KIAA0101 gene product	0.71	4.26
	100161	D14694	Hs.77329	phosphatidylserine synthase 1	1.02	1.56
	100168	D14874	Hs.394	adrenomedullin	0.46	1.17
75	100187	D17793	Hs.78183	aldo-keto reductase family 1; member C3	1	1
	100188	D21063	Hs.57101	minichromosome maintenance deficient (S.	0.97	1.4
	100217	D26600	Hs.89545	proteasome (prosome; macropain) subunit;	1.13	1.9
	100220	D28364		""Human mRNA for annexin II, 5'UTR (seq	1.11	1.53
	100287	D43950	Hs.1600	chaperonin containing TCP1; subunit 5 (e	1.13	2.09
80	100297	D49489	Hs.182429	protein disulfide isomerase-related prot	0.92	1.78
	100330	D55716	Hs.77152	minichromosome maintenance deficient (S.	1.07	1.61
	100355	D78129		""Homo sapiens mRNA for squalene epoxid	0.96	1.87
	100364	D78586	Hs.154868	carbamoyl-phosphate synthetase 2; aspart	1.49	2.46
	100368	D79987	Hs.153479	extra spindle poles; S. cerevisiae; homo	0.59	1.32
85	100398	D84557	Hs.155462	minichromosome maintenance deficient (mi	1.08	1.9
	100438	D87448	Hs.91417	topoisomerase (DNA) II binding protein	1	2.15



	100455	D87953	Hs.75789	N-myc downstream regulated	0.91	1.48
	100491	HG1153-HT1153		Nucleoside Diphosphate Kinase Nm23-H2s	0.99	1.41
	100518	HG174-HT174		Desmoplakin I	1.28	3.17
5	100528	HG1828-HT1857		""Nexin, Glia-Derived""	0.68	1.9
	100661	HG2874-HT3018		Ribosomal Protein L39 Homolog	1.1	5.44
	100667	HG2981-HT3127		""Epican, Alt. Splice 11""	0.8	1.97
	100830	HG4074-HT4344		Rad2	1.01	2.12
	101061	K03515	Hs.944	glucose phosphate isomerase	0.91	1.79
10	101131	L10838	Hs.167460	splicing factor; arginine/serine-rich 3	1.23	1.87
	101162	L14595	Hs.174203	solute carrier family 1 (glutamate/neutr	1.35	2.73
	101181	L19686	Hs.73798	macrophage migration inhibitory factor (	1.03	1.78
	101183	L19779	Hs.795	H2A histone family; member O	0.57	1.3
	101216	L25876	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	0.7	2.2
15	101228	L27706	Hs.82916	chaperonin containing TCP1; subunit 6A (	0.99	1.99
	101233	L29008	Hs.878	sorbitol dehydrogenase	0.82	2.11
	101247	L33801	Hs.78802	glycogen synthase kinase 3 beta	1.2	1.91
	101332	L47276		""Homo sapiens (cell line HL-6) alpha t	0.69	2.78
	101342	L76191	Hs.182018	interleukin-1 receptor-associated kinase	1.04	1.84
20	101396	M15796	Hs.78996	proliferating cell nuclear antigen	0.95	3.55
	101423	M18391	Hs.89839	EphA1	1	1.5
	101445	M21259	Hs.1066	small nuclear ribonucleoprotein polypept	1.21	1.96
	101505	M27396	Hs.75692	asparagine synthetase	0.93	1.6
	101525	M29536	Hs.12163	eukaryotic translation initiation factor	1.19	1.93
25	101535	M30448	Hs.251669	casein kinase 2; beta polypeptide	0.96	1.42
	101607	M38690	Hs.1244	CD9 antigen (p24)	1.11	1.25
	101624	M55998		""Human alpha-1 collagen type I gene, 3	1.17	1.98
	101758	M77836	Hs.79217	pyrroline-5-carboxylate reductase 1	1.77	3.45
	101839	M93036	Hs.692	membrane component; chromosomal 4; surfa	0.71	1.45
	101853	M94362	Hs.76084	lamin B2	0.84	1.19
30	101977	S83364		""putative Rab5-interacting protein (cl	0.89	1.9
	101992	U01038	Hs.77597	polo (Drosophila)-like kinase	0.66	1.46
	102009	U02680	Hs.82643	protein tyrosine kinase 9	1.23	3.35
	102012	U03057	Hs.118400	singed (Drosophila)-like (sea urchin fas	0.85	1.88
35	102039	U05861	Hs.201967	aldo-keto reductase family 1; member C1	0.93	2.32
	102123	U14518	Hs.1594	centromere protein A (17kD)	1	4.28
	102130	U15009	Hs.1575	small nuclear ribonucleoprotein D3 polyp	0.89	1.42
	102148	U16954	Hs.75823	ALL1-fused gene from chromosome 1q	0.8	2.95
	102210	U23028	Hs.2437	eukaryotic translation initiation factor	1.01	1.34
40	102220	U24389	Hs.65436	lysyl oxidase-like 1	1.15	2.34
	102260	U28386	Hs.159557	karyopherin alpha 2 (RAG cohort 1; impor	1.14	2.69
	102330	U35451	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta	1.05	1.7
	102423	U44754	Hs.179312	small nuclear RNA activating complex; po	1.14	2.99
	102455	U48705	Hs.75562	discoidin domain receptor family; member	1.05	2.01
45	102499	U51478	Hs.76941	ATPase; Na+/K+ transporting; beta 3 poly	1.27	1.92
	102522	U53347	Hs.183556	solute carrier family 1 (neutral amino a	0.84	1.31
	102590	U62136		""Homo sapiens enterocyte differenti	1.11	1.6
	102676	U72514	Hs.12045	putative protein	1.04	2.17
	102687	U73379	Hs.93002	ubiquitin carrier protein E2-C	0.86	2.28
50	102704	U76638	Hs.54089	BRCA1 associated RING domain 1	1.12	1.63
	102781	U83843		""Human HIV-1 Nef interacting protein (	0.9	1.39
	102784	U85658	Hs.61796	transcription factor AP-2 gamma (activat	0.98	2.16
	102827	U91327	Hs.6456	chaperonin containing TCP1; subunit 2 (b	0.96	1.62
	102935	X13482	Hs.80506	small nuclear ribonucleoprotein polypept	1.21	4.2
55	102972	X16662	Hs.87268	annexin A8	1.25	2.32
	102983	X17620	Hs.118638	non-metastatic cells 1; protein (NM23A)	1.03	1.83
	103023	X53793	Hs.117950	multifunctional polypeptide similar to S	1.58	5.44
	103038	X54941	Hs.77550	CDC28 protein kinase 1	1.32	3.79
	103075	X59543	Hs.2934	ribonucleotide reductase M1 polypeptide	1.11	2.58
60	103168	X68314	Hs.2704	glutathione peroxidase 2 (gastrointestin	0.75	3.05
	103185	X69910	Hs.74368	transmembrane protein (63kD); endoplasm	1.01	1.97
	103212	X73874	Hs.2393	phosphorylase kinase; alpha 1 (muscle)	0.95	1.72
	103223	X74801	Hs.1708	chaperonin containing TCP1; subunit 3 (g	0.97	1.77
	103260	X78416	Hs.3155	casein; alpha	1	1
65	103262	X78565	Hs.204133	hexabrachion (tenascin C; cytotactin)	1.23	3.09
	103330	X85373	Hs.77496	small nuclear ribonucleoprotein polypept	1.12	2.25
	103364	X90872	Hs.75854	SULT1C sulfotransferase	2.85	4.62
	103375	X91868	Hs.54416	sine oculis homeobox (Drosophila) homolo	1	2.48
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	1	1.53
70	103404	X95586	Hs.78596	proteasome (prosome; macropain) subunit;	0.92	1.53
	103437	X98260	Hs.82254	M-phase phosphoprotein 11	0.92	1.54
	103448	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	0.55	0.96
	103605	Z35402	Hs.194657	cadherin 1; E-cadherin (epithelial)	1.32	2.51
	103646	Z68228	Hs.2340	junction plakoglobin	0.88	1.28
75	103658	Z74615	Hs.172928	collagen; type I; alpha 1	1.06	2.98
	103774	AA092898	Hs.92918	ESTs; Weakly similar to R07G3.8 [C.elega	1.88	4.66
	104261	AF008442	Hs.5409	RNA polymerase I subunit	0.87	2.17
	104276	C02193	Hs.85222	ESTs; Weakly similar to R27090_2 [H.sapi	1.4	2.49
	104289	C16281	Hs.75478	KIAA0956 protein	1.15	1.68
80	104434	L02870	Hs.1640	collagen; type VII; alpha 1 (epidermolys	1.04	1.49
	104453	M19169	Hs.123114	cystatin SN	0.38	0.76
	104611	R98280	Hs.125845	ribulose-5-phosphate-3-epimerase	1.08	2.25
	104758	AA024661	Hs.7010	ESTs; Weakly similar to ACYL-COA DEHYDRO	1.14	1.65
	105114	AA156532	Hs.11801	adenosine A2b receptor pseudogene	0.91	1.38
85	105132	AA159501	Hs.247280	HBV associated factor	1.08	1.7
	105174	AA186613	Hs.34744	ESTs	0.95	2.05



	105280	AA232215	Hs.14600	ESTs	1	1.4
	105344	AA235303	Hs.8645	ESTs	0.72	2.02
	105516	AA257971	Hs.21214	ESTs	1.35	3.56
5	105621	AA280865	Hs.6375	Homo sapiens mRNA; cDNA DKFZp564K0222 (f	1.23	1.82
	105698	AA287393	Hs.15202	ESTs; Weakly similar to oligodendrocyte-	0.98	1.28
	105705	AA290767	Hs.101282	Homo sapiens mRNA; cDNA DKFZp434B102 (fr	0.92	1.32
	105724	AA292098	Hs.22934	ESTs; Weakly similar to ZINC FINGER PROT	0.99	1.41
	105782	AA350215	Hs.21580	ESTs	1	1
10	105799	AA372018	Hs.24743	ESTs	1.08	1.78
	105807	AA393803	Hs.16869	ESTs; Moderately similar to COLLAGEN ALP	0.95	1.34
	105891	AA400768	Hs.26662	ESTs; Weakly similar to tumor necrosis f	0.87	2.25
	105936	AA404338		ESTs	1.14	1.46
	106069	AA417741	Hs.29899	ESTs; Weakly similar to ZINC FINGER PROT	1	1
15	106103	AA421104	Hs.12094	ESTs	1.04	1.44
	106140	AA424524	Hs.14912	KIAA0286 protein	1.23	2.11
	106149	AA424881	Hs.256301	ESTs	0.83	1.48
	106154	AA425304	Hs.6994	ESTs	0.77	2.05
	106182	AA426609	Hs.10862	ESTs	0.74	2.23
20	106220	AA428582	Hs.32196	ESTs; Moderately similar to metargidin p	0.97	1.99
	106228	AA429290	Hs.17719	ESTs	0.99	1.54
	106318	AA436570	Hs.9605	pre-mRNA cleavage factor Im (25kD)	0.95	2.09
	106341	AA441798	Hs.5243	ESTs; Moderately similar to pL2 hypothe	0.98	2.66
	106432	AA448850	Hs.17138	ESTs	0.95	1.93
25	106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	1	1
	106483	AA451676	Hs.30299	IGF-II mRNA-binding protein 2	1.4	2.29
	106599	AA457235	Hs.12842	ESTs; Moderately similar to non-function	1	1.82
	106611	AA458904	Hs.26267	ESTs; Weakly similar to torsinA [H.sapie	1.49	2.78
	106654	AA460449	Hs.3784	ESTs; Highly similar to phosphoserine am	1	1.4
30	107076	AA609145	Hs.21143	ESTs; Weakly similar to fos39554_1 [H.sa	1.11	1.49
	107115	AA610108	Hs.27693	ESTs; Highly similar to CGI-124 protein	1	1.03
	107129	AA620553	Hs.4756	flap structure-specific endonuclease 1	1.13	3.63
	107159	AA621340	Hs.10600	ESTs; Weakly similar to ORF YKR081c [S.c	1.05	2.09
	107444	W28391	Hs.5181	proliferation-associated 2G4; 38kD	1.18	1.9
35	107481	W58247	Hs.27437	Homo sapiens kinesin superfamily motor K	0.99	2.74
	107516	X56597	Hs.99853	fibrillarin	0.94	1.77
	107529	Y12065	Hs.5092	nucleolar protein (KKE/D repeat)	1.05	2.29
	107531	Y13936	Hs.17883	protein phosphatase 1G (formerly 2C); ma	1.06	1.62
	107801	AA019433	Hs.173100	ESTs	1.03	1.4
40	107957	AA031948	Hs.57548	ESTs	0.95	1.46
	108565	AA085342	Hs.1526	ATPase; Ca++ transporting; cardiac muscl	0.59	1.35
	108780	AA128561	Hs.117938	collagen; type XVII; alpha 1	1	7.63
	108828	AA131584	Hs.71435	DKFZP564O0463 protein	1.33	2.56
	109060	AA160879	Hs.241551	chloride channel; calcium activated; fam	0.67	1.42
45	109112	AA169379	Hs.72865	ESTs	1.03	2.31
	109344	AA213696	Hs.86559	poly(A)-binding protein-like 1	0.97	1.55
	109412	AA227145	Hs.209473	ESTs; Weakly similar to REGULATOR OF MIT	0.76	1.87
	110780	N23174	Hs.22891	solute carrier family 7 (cationic amino	0.9	0.95
	110958	N50550	Hs.24587	signal transduction protein (SH3 contain	1.17	2.26
50	111018	N54067	Hs.3628	mitogen-activated protein kinase kinase	1.21	1.85
	111337	N79612	Hs.16607	ESTs; Highly similar to Myosin heavy cha	1	1.45
	112305	R54822	Hs.26244	ESTs	1	1
	112401	R61279	Hs.237536	ESTs; Weakly similar to F25B5.3 [C.elega	1.24	1.64
	112853	T02843	Hs.4351	EST	1.56	1.96
55	112869	T03313	Hs.4747	dyskeratosis congenita 1; dyskerin	1.03	1.57
	112992	T23513	Hs.7147	ESTs	1	1
	113048	T25895	Hs.184008	ESTs; Weakly similar to RNA-binding prot	1.37	2.26
	113063	T32438	Hs.5027	ESTs	1	1
	113179	T55182	Hs.152571	ESTs; Highly similar to IGF-II mRNA-bind	1.33	2.7
60	113573	T91166	Hs.15990	ESTs	0.76	1.47
	113811	W44928	Hs.4878	ESTs	0.79	1.51
	114086	Z38266	Hs.12770	Homo sapiens PAC clone DJ0777O23 from 7p	0.9	1.34
	114587	AA070827	Hs.180320	ESTs; Weakly similar to GOLGI 4-TRANSMEM	1.02	1.76
	114846	AA234929	Hs.44343	ESTs	1.32	2.36
65	114964	AA243873	Hs.82184	ring finger protein 3	1.1	1.84
	115047	AA252627	Hs.22554	homeo box B5	1.01	2.36
	115166	AA258409	Hs.198907	myelin protein zero-like 1	1.05	2.31
	115167	AA258421	Hs.43728	hypothetical protein	1.52	2.52
	115239	AA278650	Hs.73291	ESTs; Weakly similar to similar to the b	0.7	2.57
70	115278	AA279757	Hs.67466	ESTs; Weakly similar to BACN32G11.d [D.m	1.14	2.12
	115652	AA405098	Hs.38178	ESTs	0.82	4.67
	115875	AA433943	Hs.43946	ESTs; Weakly similar to Weak similarity	1.2	1.98
	116004	AA449122	Hs.76086	ESTs; Highly similar to small zinc finge	0.96	1.31
	116121	AA459254	Hs.48855	ESTs	0.97	1.55
75	116129	AA459956	Hs.49163	ESTs; Highly similar to putative ribonuc	1.08	2.73
	116190	AA464963	Hs.67776	ESTs	0.8	1.57
	116312	AA490494	Hs.65403	ESTs	1.37	2.65
	116732	F13779	Hs.165909	ESTs	0.92	1.8
	117602	N35020	Hs.44685	ESTs; Weakly similar to GOLIATH PROTEIN	1.15	1.84
80	117950	N51394	Hs.75478	KIAA0956 protein	1.04	2.36
	117992	N52000	Hs.172089	Homo sapiens mRNA; cDNA DKFZp586B0222 (f	0.62	1.29
	118785	N75386	Hs.111867	GLI-Kruppel family member GLI2	1	1
	119717	W69134	Hs.57987	ESTs	1	1.4
	119814	W74069	Hs.58350	ESTs	0.78	1.77
85	120128	Z38499	Hs.91448	MKP-1 like protein tyrosine phosphatase	0.86	1.46
	120242	Z98443	Hs.86366	ESTs	0.83	2.01



	120483	AA252994	Hs.1578	apoptosis inhibitor 4 (survivin)	0.74	1.64
	121054	AA398604	Hs.97387	ESTs	1.05	1.93
	121326	AA404246	Hs.97031	ESTs; Weakly similar to Similar to phyto	0.98	1.3
5	121376	AA405699	Hs.166232	ESTs; Moderately similar to SODIUM- AND	0.91	1.83
	121457	AA411448	Hs.208985	ESTs	0.91	1.59
	121780	AA422086	Hs.124660	ESTs	0.46	0.55
	121781	AA422150	Hs.98370	cytochrome P540 family member predicted	1.07	1.54
	121844	AA425732	Hs.98485	gap junction protein; beta 2; 26kD (conn	0.94	1.4
10	122059	AA431737	Hs.98749	EST	1.93	2.33
	122338	AA443311	Hs.98998	ESTs	1	1
	122354	AA443772	Hs.186692	ESTs	0.88	1.39
	122591	AA453265	Hs.99311	ESTs; Weakly similar to MRJ [H.sapiens]	2.28	2.93
	122790	AA460156	Hs.99556	ESTs	0.88	1.3
15	123398	AA521265	Hs.105514	ESTs	1	1.93
	123518	AA608531	Hs.170313	ESTs	1	1
	123673	AA609471	Hs.112712	ESTs	1	1.15
	124000	D57317	Hs.74861	activated RNA polymerase II transcriptio	0.74	1.12
	124367	N24006	Hs.99348	distal-less homeo box 5	0.67	1.1
20	124447	N48000	Hs.140945	Homo sapiens mRNA; cDNA DKFZp586L141 (fr	1.19	1.7
	125756	W25498	Hs.81634	ATP synthase; H+ transporting; mitochond	0.93	1.59
	125769	A1382972	Hs.82128	5T4 oncofetal trophoblast glycoprotein	1.65	6.76
	125852	H09290	Hs.76550	Homo sapiens mRNA; cDNA DKFZp564B1264 (f	0.72	2.26
	125924	AA526849	Hs.82109	syndecan 1	1.22	2.25
25	126037	M85772	Hs.6066	KIAA1112 protein	1.36	1.63
	126214	N29455	Hs.74316	desmoplakin (DPI; DPII)	1.93	3.55
	126414	N78770	Hs.223439	ESTs	1.21	1.66
	126737	AA488132	Hs.62741	ESTs	1	1
	126743	AA179253	Hs.172182	poly(A)-binding protein; cytoplasmic 1	1.3	2.16
30	126926	AA179546	Hs.832	ESTs; Highly similar to INTEGRIN BETA-8	2.53	2.8
	127432	AA501734	Hs.170311	heterogeneous nuclear ribonucleoprotein	1.57	2.12
	128218	H02682	Hs.99189	ESTs; Moderately similar to recombina	1.24	2.09
	128527	M31523	Hs.101047	transcription factor 3 (E2A immunoglobul	1.08	1.78
	128568	X60673	Hs.247568	adenylate kinase 3	1.23	3.48
35	128584	M11433	Hs.101850	retinol-binding protein 1; cellular	0.87	2.42
	128628	C14037	Hs.251978	EST	1.22	1.9
	128691	W27939	Hs.103834	ESTs	1.1	1.73
	128714	V00599	Hs.179661	Homo sapiens clone 24703 beta-tubulin mR	0.92	1.17
	128733	AA328993	Hs.104558	ESTs	1.34	1.94
40	128781	X85372	Hs.105465	small nuclear ribonucleoprotein polypept	0.9	1.34
	129052	AA496297	Hs.182740	ribosomal protein S11	2.59	3.19
	129095	L12350	Hs.108623	thrombospondin 2	1.04	3.2
	129241	AA435665	Hs.109706	ESTs; Moderately similar to HN1 [M.muscu	0.95	1.61
	129665	M88458	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic retic	1.28	2.63
45	129703	AA401348	Hs.179999	ESTs	0.97	1.63
	129720	AA476582	Hs.12152	ESTs; Moderately similar to SIGNAL RECOG	1.09	1.79
	129850	N20593	Hs.56845	GDP dissociation inhibitor 2	0.74	1.68
	129896	AA043021	Hs.13225	UDP-Gal:betaGlcNAc beta 1,4- galactosylt	1.43	4.19
	130069	AA055896	Hs.146428	collagen; type V; alpha 1	1.17	1.98
50	130405	H88359	Hs.155396	nuclear factor (erythroid-derived 2)-lik	1.26	1.79
	130541	X05608	Hs.211584	neurofilament; light polypeptide (68kD)	1	1
	130599	M91670	Hs.174070	ubiquitin carrier protein	1.07	1.66
	130867	J04093	Hs.2056	UDP glycosyltransferase 1	1	4.8
	131009	AA063596	Hs.22142	ESTs; Weakly similar to NADH-CYTOCHROME	0.93	1.05
55	131028	U20240	Hs.2227	CCAAT/enhancer binding protein (C/EBP);	1	1.23
	131083	U66661	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.1	1.8
	131091	T35341	Hs.22880	ESTs; Highly similar to dipeptidyl pepti	1.28	1.98
	131144	C14412	Hs.23528	ESTs; Highly similar to HSPC038 protein	1.43	2.06
	131148	C00038	Hs.23579	ESTs	0.88	3.38
60	131164	Y00503	Hs.182265	keratin 19	1.19	2.77
	131185	M25753	Hs.23960	cyclin B1	0.86	3.84
	131219	C00476	Hs.24395	small inducible cytokine subfamily B (Cy	0.66	2.96
	131454	AA455896	Hs.2699	glypican 1	0.99	1.54
	131687	L11066	Hs.3069	heat shock 70kD protein 9B (mortalin-2)	1	1.18
65	131689	AA599653	Hs.30696	transcription factor-like 5 (basic helix	1	1.95
	131692	D50914	Hs.30736	KIAA0124 protein	1.55	2.39
	131786	AA135554	Hs.32125	ESTs	1	1.33
	131843	AA195893	Hs.184062	ESTs; Moderately similar to putative Rab	0.83	1.63
	131860	U02082	Hs.334	Oncogene TIM	1.08	2.2
70	131884	H90124	Hs.3463	ribosomal protein S23	1.23	1.24
	131903	AA481723	Hs.3436	deleted in oral cancer (mouse; homolog)	0.91	1.18
	131945	M87339	Hs.35120	replication factor C (activator 1) 4 (37	1	2.8
	131958	AA093998	Hs.3566	ESTs; Highly similar to phosphorylation	0.87	1.36
	131964	W42508	Hs.3593	ESTs	1	1.25
75	132001	J00277	Hs.37003	v-Ha-ras Harvey rat sarcoma viral oncoge	1.12	1.43
	132040	AA146843	Hs.172894	BH3 interacting domain death agonist	1	1.55
	132065	D82226	Hs.211594	proteasome (prosome; macropain) 26S subu	0.89	1.27
	132109	AA599801	Hs.40098	ESTs	1	1.05
	132112	AA150661	Hs.40154	jumonji (mouse) homolog	0.99	1.44
80	132123	AA447123	Hs.250705	ESTs	1.06	2.46
	132162	H89551	Hs.41241	ESTs	1.08	2.46
	132180	AA405569	Hs.418	fibroblast activation protein; alpha; se	1.02	4.56
	132309	AA460917	Hs.2780	jun D proto-oncogene	1.16	1.8
	132371	AA235448	Hs.46677	ESTs	0.8	1.26
85	132618	AA253330	Hs.5344	adaptor-related protein complex 1; gamma	0.5	1.49
	132736	U68019	Hs.211578	MAD (mothers against decapentaplegic; Dr	1.21	1.81



	132771	AA488432	Hs.56407	phosphoserine phosphatase	1	1.3
	132833	U78525	Hs.57783	eukaryotic translation initiation factor	0.91	1.43
	132922	T23641	Hs.6066	KIAA1112 protein	1.16	1.53
5	132959	AA028103	Hs.61472	ESTs; Weakly similar to unknown [S.cerev	1.02	1.88
	132994	AA505133	Hs.7594	solute carrier family 2 (facilitated glu	0.72	2.97
	133005	C21400	Hs.103329	KIAA0970 protein	0.88	1.34
	133065	X62535	Hs.172690	diacylglycerol kinase; alpha (80kD)	0.93	1.23
	133083	N70633	Hs.6456	chaperonin containing TCP1; subunit 2 (b	1.14	1.76
10	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso	0.97	1.43
	133134	T89703	Hs.65648	RNA binding motif protein 8	1.1	1.8
	133195	AA350744	Hs.181409	KIAA1007 protein	2.29	2.69
	133313	AA249427	Hs.70704	ESTs	1.07	1.68
	133331	T62039	Hs.158675	ribosomal protein L14	0.85	1.18
15	133438	D13370	Hs.73722	APEX nuclease (multifunctional DNA repai	0.91	1.45
	133445	T99303	Hs.73797	guanine nucleotide binding protein (G pr	0.94	1.68
	133483	X52426	Hs.74070	keratin 13	0.85	1.14
	133492	L40397	Hs.74137	transmembrane trafficking protein	1.1	1.69
	133504	W95070	Hs.74316	desmoplakin (DPI; DPII)	0.7	6.21
20	133517	X52947	Hs.74471	gap junction protein; alpha 1; 43kD (con	0.95	1.3
	133540	D78151	Hs.74619	proteasome (prosome; macropain) 26S subu	0.91	1.25
	133594	L07758	Hs.172589	nuclear phosphoprotein similar to S. cer	0.84	1.29
	133627	U09587	Hs.75280	glycyl-tRNA synthetase	1.09	1.99
	133671	T25747	Hs.75471	zinc finger protein 146	1.02	1.5
25	133859	U86782	Hs.178761	26S proteasome-associated pad1 homolog	1.11	3.33
	133865	F09315	Hs.170290	discs; large (Drosophila) homolog 5	1.84	6.7
	133913	W84712	Hs.7753	calumenin	1.15	1.86
	133963	L34587	Hs.184693	transcription elongation factor B (SIII)	1.3	1.91
	133982	U47621	Hs.207251	nucleolar autoantigen (55kD) similar to	1.3	1.99
30	134100	L07540	Hs.171075	replication factor C (activator 1) 5 (36	0.72	1.65
	134110	U41060	Hs.79136	LIV-1 protein; estrogen regulated	1.04	1.62
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	1	1.55
	134161	U97188	Hs.79440	IGF-II mRNA-binding protein 3	0.82	1.95
	134193	F09570	Hs.7980	ESTs	0.98	1.48
35	134367	X54199	Hs.82285	phosphoribosylglycinamide formyltransfer	1	2.8
	134402	U25165	Hs.82712	fragile X mental retardation; autosomal	1.26	2
	134457	D86963	Hs.174044	dishevelled 3 (homologous to Drosophila	1	1.47
	134469	X17567	Hs.83753	small nuclear ribonucleoprotein polypept	0.94	1.57
	134498	M63180	Hs.84131	threonyl-tRNA synthetase	1.2	2.64
40	134501	W84870	Hs.211568	eukaryotic translation initiation factor	0.84	1.36
	134507	M63488	Hs.84318	replication protein A1 (70kD)	1.7	2.93
	134548	U41515	Hs.85215	Deleted in split-hand/split-foot 1 regio	1.46	2.73
	134599	X99226	Hs.86297	Fanconi anemia; complementation group A	1.36	2.22
	134692	R73567	Hs.8850	a disintegrin and metalloproteinase doma	0.77	1.64
45	134693	N70361	Hs.8854	ESTs	1.09	1.82
	134806	Z49099	Hs.89718	spermine synthase	0.98	1.35
	134821	Z34974	Hs.198382	plakophilin 1 (ectodermal dysplasia/skin	0.99	1.4
	134864	Y08999	Hs.90370	actin related protein 2/3 complex; subun	0.95	1.42
	134914	U29615	Hs.91093	chitinase 1 (chitotriosidase)	1.16	1.29
50	134953	L10678	Hs.91747	profilin 2	0.95	1.76
	134993	AA282343	Hs.9242	purine-rich element binding protein B	0.98	1.73
	135051	C15324	Hs.93668	ESTs	1.35	2.11
	135158	U51711		Human desmocollin-2 mRNA; 3' UTR	0.86	1.16

55 Table 1B shows the accession numbers for those pkeys in Table 1A lacking unigenelD's. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the Accession column.

60 Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT	Accessions
65	100661 23182_1	BE623001 L05096 AA383604 AW966416 N53295 AA460213 AW571519 AA603655
	100667 26401_3	L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700 AW606203 BE069721 AW382138 AW803776 BE463954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 AI632506 T29066 AI783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153 BE546944 T69231 AW377441 AA907406 H50799 AW051416 AI420712 BE620922 AI279161 AA992549 W47198 BE005241 AI342696 H50700 70 AI969974 AI863855 AA374490 AW130675 AI950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 AI673068 AI887890 AW804171 AI675961 AW804172 AA778841 AL048050 AI127757 AI095568 AW204965 AW468978 W31898 AI052595 AI278771 BE464018 AI081503 AI824196 AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 AI280942 T27619 BE621435 N66010 AW589527 AI160414 AA283090 AA962536 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE565671 C00444 AA054555
75	100668 26401_3	L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700 AW606203 BE069721 AW382138 AW803776 BE463954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 AI632506 T29066 AI783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153 BE546944 T69231 AW377441 AA907406 H50799 AW051416 AI420712 BE620922 AI279161 AA992549 W47198 BE005241 AI342696 H50700 80 AI969974 AI863855 AA374490 AW130675 AI950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 AI673068 AI887890 AW804171 AI675961 AW804172 AA778841 AL048050 AI127757 AI095568 AW204965 AW468978 W31898 AI052595 AI278771 BE464018 AI081503 AI824196 AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 AI280942 T27619 BE621435 N66010 AW589527 AI160414 AA283090 AA962536 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE565671 C00444 AA054555
85	101332 25130_1	J04088 NM_001067 AF071747 AJ011741 N85424 AL042407 AA218572 BE296748 BE083981 AL040877 AW499918 AW675045 H17813 BE081283 AA670403 AW504327 BE094229 AA104024 AI471482 AI970337 AA737616 AI827444 AW003286 AI742333 AI344044 AI765634



5			AI948838 AW235336 AW172827 AA095289 BE046383 AI734240 W16699 AI660329 AI289433 AA933778 AW469242 AA468838 AA806983 AA625873 W78031 BE206307 AA550803 AI743147 AI990075 AA948274 AA129533 AI635399 AA605313 AI624669 AW594319 AI221834 AI337434 AA307706 BE550282 AI760467 AI630636 AI221521 AW674314 AW078889 AI933732 AI686969 AI186928 AW074595 AI127486 AL079644 AI910815 H17814 AA310903 AW137854 T19279 AA026682 AA306035 AW383390 AW383389 AW383422 AW383427 AW383395 H09977 AA306247 AA352501 AW403639 F05421 AA224473 AA305321 H93904 AA089612 AW391543 AW402915 AW173382 AW402701 AW403113 R94438 N73126 H93466 AA090928 AA095051 T29025 AW951071 L47277 L47276 AI375913 BE384156 W24652 AA746288 AA568223 BE090591 H93033 N57027 AA504348 AA327653 AW959913 N53767 AA843715 AI453437 AW263710 AI076594 AA583483 AW873194 AW575166 AI128799 AI803319 AL042776 AW074313 AI887722 AI032284 AA447521 AI123885 N29334 AI354911 AW090687 AA236763 AA435535 AA236910 AA047124 AA236734 AW514610 H93467 AA962007 AI446783 AA127259 AI613495 AI686720 AI587374 AA936731 AA702453 AI859757 AA216786 AI251819 AI469227 AA806022 AI092324 N71868 AA968782 AA236919 AA809450 AA227220 AA765284 AI192007 AA768810 AA805794 AA729280 AA806238 AW768817 N71879 AI050686 AA505822 AA668974 AI688160 BE045915 AW466315 AA731314 AA649568 AA834316 AW591901 AW063876 AW294770 AI300266 AI336094 AI560380 AA721755 H09978 D20305 D29155 AW821790 BE150864 F01675 AI457474 AW466316 AA550969 AA630788
10			BE561958 BE561728 BE397612 BE514391 BE269037 BE514207 BE562381 BE514256 BE514403 BE514250 BE397832 BE269598 BE559865 BE396881 BE560031 BE514199 BE560037 BE560454 AC004770 W05005 AA356068 AA094281 H29358 T56781 AW875313 L37374 BE312466 BE311755 BE207106 BE293320 BE018115 AW239090 BE548830 AW247547 AA776062 BE397382 AA486713 T10111 T09340 AW498981 BE547280 AA356003 AW581520 AW875331 AA580720 AW875336 BE276873 BE408229 AW188148 BE255166 BE253761 AW793727 AW373141 AW581548 AA471223 AA305950 BE263976 AA626820 BE257409 AW360962 AA090655 C00312 BE312741 BE407213 AA209352 AW298199 AW248553 AW297794 AW731722 BE300586 AW731972 AW615446 BE301599 AW615520 AA486714 AW440257 AA196516 AA564630 AA618079 AW192592 AW474985 AA604580 AI627461 AA765440 AI680394 AL135548 AI683224 AI581126 AW245096 AW194154 H29274 N70363 AA629758 AA580602 AA862006 AI863841 AI097667 AI928583 AI358774 BE243487 AA620553 AA653297 AA292690 T10110 Z38906 AA908544 AA340930 AI185438 T03328 T28844 AI687010 AI864965 AI872575 BE388740 T56780 AW373138 BE258717 AA699671
15	100780	458_127	AU076916 BE298110 AW239395 AW672700 NM_003875 U10860 AW651755 BE297958 C03806 AI795876 AA644165 T36030 AW392852 AA446421 AW881866 AI469428 BE548103 T96204 R94457 N78225 AI564549 AW004984 AW780423 AW675448 AW087890 AA971454 AA305698 AA879433 AA535069 AI394371 AA928053 AI378367 N59764 AI364000 AI431285 T81090 AW674657 AW674987 AA897396 AW673412 BE063175 AW674408 AI202011 R00723 AI753769 AI460161 AW079585 AW275744 AI873729 D25791 BE537646 T81139 R00722
20	100830	4002_1	J04129 NM_002571 AA293088 AA477016 AA404631 T28299 AA476904 AA433965 AA430486 AA495907 AI151391 AA291495 AA402723 W25651 AA706816 AI826712 AW296294 AA293479 AI276581 AW044154 AI080180 AI417985 AI274168 AI474212 AA495908 AA635664 AI092114 AI804952 AA479874 AI597661 AI420511 AA479738 AA421417 AA421247 AA436220 AL047797 M34046 N42277 AA228076 W02698 AI420297 AA434011 AI369971 AA479731 AI865541 AI418020 AA421246 AA452764 AL048051
25	100906	4312_1	NM_006769 U24576 AW161961 AW160473 AW160465 AW160472 AW161069 AI824831 AW162635 AI990356 AW162477 AW162571 AI520836 AW162352 AW162351 AW162752 AI962216 AI537346 AA853902 H17667 BE045346 BE559802 BE255391 AA985217 AA235051 AI129757 AW366451 T34489 D56106 D56351 AI936579 AW023219 AW889335 AW889120 AW889232 AW889175 BE093702 AW889349 AA147546 AI952998 AA912579 AI143356 AW902211 R64717 AW157236 AI815242 D45274 AW263991 AA442920 AA129965 AL035713 AI923255 AI949082 AI142826 AI684160 AI701987 AI678954 AI827349 BE463635 AW628092 AW302281 AA493203 BE348856 BE536419 AW193969 AW673561 AW592609 AI224044 H43943 AA091912 R49632 R48353 AI568409 R48256 AI198046 H27986 H43899 AI678759 AI680310 AI624220 H17052 AA156410 N56062 AI699430 AA664529 T09406 T10459 AA627506 AI379584 N83831 N88633 AW022651 AA971281 AA248036 AI039197 AI914689 AA973825 AL047305 AA129966 AI798369 AW264348 AI445879 AI658759 N67924 AI933507 AI216121 AI333174 T10972 AI375028 AI186756 AI273778 AA610487 AI797946 AA853903 AA903939 AI338587 AI278494 AW627595 AA904019
30	100930	16865_1	M86849 AA315280 NM_004004 AA315269 BE142653 AA461400 AW802042 BE152893 AW383155 AA490688 AW117930 AW384563 AW384544 AW384566 AW378307 AW378323 AW839085 AA257102 AW378317 AW276060 AW271245 AW378298 AW384497 AI598114 AW264544 AI018136 AW021810 AA961504 AW086214 AW771489 AW192483 AI290266 AW192488 AW384490 AW007451 AW890895 AA554460 AA613715 AW020066 AI783695 AI589498 AI917637 AW264471 AW384491 AI816732 AW368530 AW368521 AW368463 AA461087 AI341438 AI970613 AI040737 AI418400 AA947181 AA962716 AI280695 AW769275 AW023591 AI160977 AA055400 N71882 AA490466 AW243772 AW316636 AI076554 AW511702 N69323 H88912 AA257017 AI952506 H88913 AI912481 AA600714 BE465701 N64149 C00523 N64240 AA677120 R61573 BE005029 X98091 AA297307 BE537267 BE566138 BE566139 F11561 BE564795 BE568776 AW064005 BE566479 BE380035 BE567012 BE568634 BE566568 AA298060 BE566043 BE568813 BE568618 AA283070 BE565414 BE566738 BE568585 BE565667 BE566116 BE566433 U62136 AF049140 BE567057 BE567297 BE567403 BE564316 BE567400 BE568854 BE566588 AA448772 AA071363 AW732642 BE564996 AA297763 AA278550 AA421083 AA298184 AA091007 AA984577 AA205916 N28759 AL031291 C15757 C15761 H02728 BE566410 AA129335 AA419499 N87741 BE379689 BE004824 BE379611 D25874 AA148454 AA323654 AW950311 AA448795 AW749423 AA773386 AA773843 AW020327 BE348580 BE504258 BE549990 BE220200 AI673334 AI202679 AA975515 D61421 AI168688 AA102843 AW246621 AI276203 AI074054 AI633824 AI962927 AI148926 N50969 AI308911 AA410994 AW373025 AA148455 H02620 AA688293 AI246318 N22220 AI917777 AI050943 AI097286 AA663794 AW368662 AW627826 AW078734 AI253060 AA749154 AA832236 AI192358 AW024676 AA448676 AA764891 BE439467 AA661534 AA258061 AI090546 AA995157 AI051011 AA584421 AI026032 AW591338 AW589563 AA776914 AW024684 AA421002 F09219 BE464500 AI383595 AA954244 AA601583 AA737304 AA195549 AA805778 AI055876 AA164942 AW013961 AI672608 AW514211 D59441 AW582574 AA160935 BE566501 BE564612 BE565353 BE568195 BE565447 BE568302 BE566097 BE565470 BE564249 AL036217 AW749424 BE567494 AA102842 AA314761 AV661237 C14211 AA651866 AW798997 AA470605
35	102221	3861_1	AF112213 AL050318 T24804 AW248136 BE386341 BE263177 W16677 BE250224 BE563669 BE267405 BE546577 AV651354 AV651292 AI346903 AI539128 AI189171 S83364 AW073849 AI816760 AW073309 AI422690 AA296692 AI860301 AI805446 N77735 AI340328 BE092530 AW028742 BE088442 AA657742 AA742438 AW170086 AI038920 AI432379 N36073 AI936194 AA868655 AA983612 AI077505 BE080433 AI375014 AI126547 AI348244 AI346077 AI748952 N26915 AI753574 AI093341 AI278762 BE092517 N74204 H06158 T58149 AI129303 N58366 AA524456 BE122661 AA542925 AI246120 AI735203 AA706829 AA877544 AI082289 AA926687 N92840 AW249798 AA934763 AW998363 AI128632 N25202 AI240209 AW118892 N80744 R35655 AI342321 AI340141 AW878792 AI857321 H09610 W04601 AW006650 AA126006 AA553675 AI052791 AW059835 AI041906 AA814658 AW002059 AA729483 AI609301 AA994633 AA903651 AI459183 T95072 AW088630 AA126112 AI800091 AI561215 H17502 AW475072 AI819003 AI683272 AI262701 AW793140 T81787 R99586 AI275160 AI310420 AI698929 AA159174 AI827968 F30305 F30309 AA806662 AI091923 AW878722 AA583430 AW571913 AI674584 AA292533 AI079471 AA642325 AA719050 AW793172 AA305476 AW103745 T23459 N79525 AI784438 AA534551 AW193751 AI074360 BE281214 T32229 W25066 W01205 T63086 AW795348 AI361287 AW795353 AW795349 AA594759 AI400295 D11489 AI370689 AA482356 AA485295 W40151 AA564661 AW300745 AI346938 AI374975 AI423782 AW193899 AA612604 AI183409 AA996156 AW366963 AW366977 AI284860 AA846503 AI985064 AA844576 AA737921 AA873274 BE241546 BE241540 AA484058 AW468970 AA127876 AA159120 AW001568 AW795213 AW795258 AW795330 BE250589 BE387572 AA910895 AA161217 BE250380 W31500 T95167 AI719306 AI359224
40	101809	32963_1	BE258778 BE281230 BE410044 T33723 AW672694 AW410439 NM_006429 AF026292 T35505 BE542333 T08940 AU076737 AW247471 BE393215 AW328640 BE542408 T32170 BE302544 T31955 BE206898 BE275738 T32570 BE386426 BE298746 BE389937 BE293991 BE315289 BE389578 R34739 R15312 BE279365 BE277756 AL036019 T33725 BE277779 BE302962 AL047294 BE276505 T09070 T33673 BE312580 AW387774 BE257175 AW674367 BE253331 BE270344 BE299831 BE273576 T32062 AI751831 BE618381 AA304899 BE252268 U46364 BE256790 BE207199 BE256209 BE251941 BE250791 BE313955 BE269806 BE543623 BE279212 BE252289 T31699 BE262220 T31669 AA315781 AA192212 N84547 BE292737 BE259631 AA232179 AI133144 T31292 AA315945 BE407301 BE251184 BE409006 AI880158 AI904003 AI904114 AW651768 AW651763 R58247 BE271897 U83843 C05298 BE261609 BE255973 AA351650 N84631 BE263637 AW452910 AA328465 AA324549 AW579525 BE252296 BE257551 AL048332 BE208630 AA359336 AW327897 AA151742 AA305816 BE076862 BE076796 BE263161 AA323785 AA676588 AA626565 AA078917 W87657 R09002 R94021 AA312032 BE276665 AA295608 AW407162 AA329374 AW877912 N27885 AA369256 AA360968 BE250476 N85427 BE265569 AI278639 AI816576 AI691037 AW328583 AI567949 AI983455 AI927732 AI811297 AI571508 AW073674 BE296039 BE467326 AI828796 AI816578 AW511604 AI921213 AW152427 AI795787 AI801618 AW168866 AI628144 AI890339 AW173690 AW511540 BE535620 AA383014 BE301164 AI866596 AW514909 AA658050 AW575243 AA074631 AI093488 AW575408 AW675443 AW615636 AW732207 AW377638 AA321784 AA641629 AA633105 AA527640 AW129146 AW615672 BE394607 AA483902 AW475032 BE378532
45	102590	15932_1	
50			
55			
60	101977	29073_1	
65			
70			
75	102781	20812_1	
80			
85			



5			AA872808 AI469388 AW105268 BE047301 AW591843 AW410066 AW517153 AI950495 AA746641 AI914878 AA873185 AI696911 AA548625 AA911505 AA148762 AW674535 AI587329 BE328328 AW270348 AA158225 AW117705 AW474997 AW519193 AA614757 AW664383 AI082647 AW590973 AI476711 AA192213 N88741 BE464552 AW072679 AI453708 AA152166 AA805924 AI581078 AI125768 AW173484 AI961980 BE300766 AI199698 AI636792 AW247333 AW272861 AA078818 AA150012 AA551232 AA678821 AW873869 AW768266 AI660315 AA319210 AA814551 AA157994 AA318886 AI582962 AW089224 AI356098 AI343694 AW072598 N21054 AI301249 AA742924 H17917 AW328584 AW248898 AI751830 AA907816 R08898 AW087989 AI828300 AA148596 AI269577 T33426 AA213571 AI973201 AA666279 R49612 AI573183 AW799762 AW410068 AW769666 AI962097 AI475204 D57490 AW517531 BE245270 AW470008 T33427 AW005731 AI795795 T23753 AW272981 T15747 AA552875 T23644 AW361289 AI758558 BE207435 AA876958 T03361 AA883569 F37533 AA582321 AW082524 R42212 AA973847 T18900 AA086202 AI559867 AI302418 AA948667 AA745670 T08939 T33724 T33722 BE621568 D57489 D25906 BE621151 F16510 C05966 T35127 AA630427 AI933481 AA309426 AI918440 BE561854 BE618866 BE394675 BE296173 AW951687 BE383739 BE616141 BE312730 BE535351 AW080575 BE313330 BE616664 AI354390 AA847315 BE544509 BE515212 BE297833 BE278808 BE544844 AW090178 AI890664 BE546708 AW189943 BE274412 BE382399 BE266392 BE254949 BE280696 BE383237 BE261756 BE257721 BE312683 BE275476 BE514880 BE545314 BE313587 BE384537 BE386691 BE264813 AW592575 AI336332 AI278641 AI795791 BE222662 AW249316 AA314361 AL036012 AW402923 BE266845 AA075945 AA314436 BE384640 AW731769 AW957077 AA552234 AA573560 AW367038 AA313399 AI983873 BE410159 BE263803 BE514339 BE409073 BE281296 BE543396 BE395387 BE088360 BE546946 BE546570 BE390626 AA074638 AA301821 AW845230 AW582379 AI949222 AW029572 AA515843 AW272394 BE250234
10	119221	102947_1	C14322 W74050 AI074232 AA595624 BE048955 AI148417 AI583145 AI473460 AI801688 AW573593 AI950741 AI628140 AW467921 R98105 AI149258 AI247584 AI078378 AI139850 AA489411 W24744 R98104 AI033826 AA699589 AI033120 N55544 W88984 AW970771 AA703362 AA099138 AA706792 AA046150 H98981 AI916674 AA953018 AI972749 AI921343 AA909044 AA094751 AI203124 AA582143 AI446654 AW235415 R70377 AA099236 F20703 AA524436 R69484 H04043 D60988 D60337 AI204246 AI204250 AI194050 H83265 T63524 AA304359 AW960551 AI672874 AI749427 AA227777 AW027055 AA971834 T49644 T54122 AI983239 AI808233 T91264 T96544 AI350945 AI709114 R72382 T48788 R48726 AW385418 AI095484 T49645 AA928653 AA570082 AW007545 T57178 AA516413 AA913118 T57112 AA564433 AA774503 AA367671 T59757
15			W78816 AI720806 AI633854 AI632086 AI668663 N70894 AW571809 AI383592 AI201348 W80715 N91880 AW963101 AA339011 AB033023 BE391906 BE275965 BE277872 BE003882 AA313774 BE019159 BE298024 BE299727 BE300011 BE390277 BE394764 N87550 BE409419 BE408652 BE408197 AL119332 AA622427 AI816265 AA610118 T07318 AA019839 AA634430 BE205794 BE049461 AI042322 AI652711 AI917645 AA630045 AW191969 AI817882 T17271 AI803663 AI095533 H46019 AW592438 AI624836 AI675552 D51149 AW132058 AA639614 AI925762 AW088153 T17455 AA018640 AW751475 BE300241 AI816255 BE391981 AW408671 AA353910 AW875446 AW875703 AW875926 AW875645 AW875647 AW938037 AL138042 AW892619 BE243018 AW995454 BE246381 BE009082 BE278921 AW967842 AA262454 H30121
20	125831 128192 113195	1522905_1 45743_3 178688_1	W72062 AF088057 W76255 AI827219 AI631461 AW449295 AI354957 AI913803 T62772 AI222040 T62921 T63781 AI678765 H12175 R14664 AI914049 AA995383 H08009 H19418 AW953728 AI358021 AA587361 AI269377 AA369905 AW957113 H27693 AI300474 H73776 W74397 AA579604 AI131018 W72331 AI719085 AA568348 AI859045 AI814819 AI888714 BE467470 AW131268 H19419 H27694 AI342165 AI914155 AA534872 BE018176 R60206 H11647 R45641 AI860466 BE301656 AI125453 AI498120 AA593735 AA879110 AI016404 T35018 AA588397 AW449767 AA470365 BE501139 AA588354 AI337500 AW078532 Z41279 AI125449 AA935725 AA404338 L42583 NM_005554 L42601 BE183076 AI541221 BE140567 L42610 V01516 J00269 AW275792 AW383052 AW380143 AI541102 BE612846 AI541344 AW238368 BE613405 BE615705 BE615530 BE615301 AW379823 AW794706 AA194806 AA194992 AW384024 AW384000 AA641239 AI246504 AI540333 AW238681 AA640939 AI540863 AI608860 AW862564 AW366725 AW368983 AW366870 AA596020 AW794721 AW794511 AI591181 BE182523 AW794644 AW794620 AI935234 AI608903 AI608623 AW797060 AW084935 BE182517 BE182319 AI890082 AW238346 AW797012 BE182522 AW794838 AI608794 AW304289 AA147193 AA595995 AW381128 AW366720 AA583718 AI828416 BE122864 AW368343 AA431080 AW082039 AW380976 AA587144 AA443636 AW872937 AW794448 AW378382 AW085761 AW794718 AW263895 AA583587 AA583991 AA583994 AA586886 AA586880 AW368365 AI814460 AA586991 AI282829 AW378406 AA586721 AI609242 AA431973 AA232959 AI831095 AW263854 AW378391 AW378415 AW378381 AA036990 AW238395 AI285446 BE208219 BE049526 AA583605 AA583918 AW366711 AI285580 AW082642 AI285712 AA582875 AW591216 AW368719 AW378408 BE122835 AA582976 BE350422 AA418328 AI541454 AI565930 AA583700 AA150575 AW238427 AI287474 AA912658 AA584223 AW238528 C17918 AW136169 AA159847 AI923797 AI609009 BE182479 AI915198 AW378114 AA147179 AA584239 AA150532 AW168862 AW085999 AW082480 AA659742 AW079703 AI872793 AA583981 AI824571 BE182316 BE182507 AA233331 AI824572 AI540586 D29492 BE182931 AA036948 BE551821 D29401 AW378365 C00141 D29181 D29567 AW103359 W95238 AI991663 AA587298 BE184608 AA099833 W95121 W95150 D29584 AI934111 D29456 D29533 AW265380 D29290 AW238463 AA121041 D29204 AA595925 D29441 AW081840 AA587018 D29323 AA582891 BE182433 BE182437 BE158295 BE182434 AW015534 AA314369 AA290715 BE568683 AW629494 D28364 AW995678
25	119861 112973	238266_1 4868_1	AI907114 AA580734 AL041945 AA101515 AA121344 D78130 NM_003129 AA341650 T84166 AF098865 AA130976 BE089553 T66122 AW175590 F05344 AI114790 R12900 AA194871 AA132298 D78129 AA132213 AW948930 AW948919 AA263053 AW946593 AW948840 AA278558 R50895 N26940 N40818 AW021255 AA054851 AA663379 AW948795 AW948893 AA400356 AW948911 N85024 W78844 AI341546 AI760182 AA286783 BE617763 BE617263 AW263690 BE049454 BE617928 AW515038 AW950584 AA601009 AI079194 AA147204 AW083163 AA130981 AI218369 AA604784 AI806257 AI559556 AA232318 AA258065 AI471982 AA687949 AI143944 N30172 AA400196 AI769049 AI084342 AI221380 AA948469 AI802469 H05720 AA113270 AA158138 AA076231 AI521024 AI810962 AI133616 AA805106 AA101516 R40052 R50778 R43280 T65036 AW131924 AA114251 AA152331 F09650 AA580614 AA558927 C75491 Z38352 AA954595 C75606 W80742
30	129402 105936	47367_1 260931_1	D56165 M36981 X58965 NM_002512 BE379177 AA314836 BE256445 BE252016 AW248343 AI720933 AW085701 BE386050 BE619742 BE277805 AA147951 AA603113 BE253293 AI246588 AI183405 AI954174 AI126891 AI829101 AI123832 AW129670 AA471268 AW170242 AW873079 AA148011 AI608620 AA482961 AI003658 H43261 AA657978 AI735072 R83138 AA722002 AA626271 AW273877 BE464626 AA071483 AA429973 AA494342 AA620436 AA775597 AA775601 AA826847 AI192585 AA826359 AA411159 AI193419 AI204013 AA705323 AA716255 AI784611 AI081144 AI128227 AA828464 AI148911 AI493446 AI626084 AI189180 AI721196 AI190618 AA284987 AI128543 AA632064 AI333073 AI278470 AA131688 AI491768 AA937581 AA630065 AA834257 AW249841 AA583742 AI309756 AA961676 AI760860 AA557818 AA954238 H43655 AI302564 AA127545 AI609219 H20426 AI042292 AI056466 AA581836 W47002 AA422057 AA937673 F29757 AA829208 AW327462 AA372098 W02144 AA036805 AA487365 AA961037 AI139946 AA487250 AA737118 AI952504 AI242293 AA650552 AI708401 AI633133 AA630848 AA654317 F24128 AI434165 W46252 AW043879 AI033763 F37228 AA687809 N49087 AA876981 AA506947 AI914572 AI833284 F22253 AA026222 R50166 AI219267 N27095 AA496512 AI784222 AI289904 AA513146 AA528547 AA418700 F36721 AI880700 AI601170 AI862851 AI708633 AA524499 AA642220 AA496628 AI718709 W80579 AI720547 F20718 AA649943 AA588229 N40503 H46029 BE262669 BE391069 BE537538 AI510751 AI906968 AI318611 H46099 AI472604 T60667 AA373087 W32479 AA514034 BE619183 AA134672 AA127544 H26942 BE536689 AW327461 AA422139 AW262357 AW327348 F33510 AI630382 AW827126 F27133 AI335189 AW517599 W80471 AA885814 N89681 BE393173 AA617760 AA584268 AA460537 AA446261 H20425 N64040 AW276801 AA316367 AA071232 BE545409 AA308292 BE274447 AA380861 AA340038 AA341806 AA865579 AI018634 AI766314 AI919302 AA872367 AA991404 AI906961 AA888375 BE621012 AA505388 AA935192 AA290828 R50220 H50814 H44721 AW951723 AA514796 AA418708 AW673377 AA379622 AA977995 AA708224 AA708216 AI318249 AI318233 AA411160 AA026221 AA316774 AA486908 AI500094 AA096362 AW583742 BE536422 BE618653 R70203 AA131732 AA345048 BE562720 T28342 NM_004415 AL031058 M77830 BE149760 AW752599 AW848723 AW376697 AW376817 AW376699 AW848371 AW376782 AW848789 AW361413 AW849074 AW997139 AW799304 AW799309 BE077020 BE077017 BE185187 AW997196 BE156621 BE179915 BE006561 BE143155 AW890985 BE002107 AW103521 AA857316 AW383133 BE011378 AW170253 BE185750 AW886475 BE160433 J05211 BE082576 BE082584 BE004047 AW607238 AW377700 AW377699 BE082526 BE082505 BE082507 BE082514 AW178000 AW177933 AI905935 AW747877 AW748114 BE148516 AW265328 AW847678 AW847688 AW365151 AW365148 AW365153 AW365156 AW365175 AW365157 AW365154 AW068840 BE005272 AW365145 BE001925 BE182166 BE144243 BE001923 AI951766 AI434518 BE184920 BE184933 AI284090 BE184941 AW804674 BE184924 C04715 W39488 AW995615 BE184948 BE159646 AW606653 AA099891 AA131128 AA337270 AA340777 AW384371 AA852212 R58704 AW366566 AW364859 AA025851 AA025852 AA455100 AA719958 AW352220 AW996245 BE165351 BE073467 AA377127 AW890264 AW609750 AW391912 AW849690 T87267 AW853812 AA852213 W74149 BE009090 AA056401 H91011 AW368529 AW390272 C18467
35			
40	129466	2094_50	
45			
50	100220 100355	45374_1 12538_1	
55		H05719 F13446	
60	100491	34803_1	
65			
70			
75	100518	13165_1	
80			
85			



5

10

15

20

25

AW674920 N57176 AA026480 AW576767 H93284 AA026863 AW177787 AA026654 AW177786 BE092134 BE092137 BE092136 AW177784  
AI022862 BE091653 AW376811 AW848592 AA040018 BE185331 BE182164 AA368564 AW951576 T29918 AA131077 W95048 W25458  
AW205789 H90899 N29754 W32490 R20904 BE167181 BE167165 N84767 H27408 H30146 AI190590 C03378 AI554403 AI205263 AA128470  
AI392926 AF139065 AW370813 AW370827 AW798417 AW798780 AW798883 AW798569 R33557 AA149190 C03029 AW177783 AA088866  
AW370829 AA247685 BE002273 AI760816 AI439101 AW879451 AI700963 AA451923 AI340326 AI590975 T48793 AI568096 AI142882 AA039975  
AI470146 AA946936 BE067737 BE067786 W19287 AA644381 AA702424 AI417612 AI306554 AI686869 AI568892 AW190555 AI571075 AI220573  
AA056527 AI471874 AI304772 AW517828 AI915596 AI627383 AI270345 AW021347 AW166807 AW105614 AI346078 AA552300 W95070  
AI494069 AI911702 AA149191 AA026864 AI830049 AI887258 AW780435 AI910434 AI819984 AI858282 AI078449 AI025932 AI860584 AI635878  
AA026047 AA703232 D12062 AW192085 AA658154 AW514597 AW591892 T87181 AA782066 AW243815 AW150038 AW268383 AW004633  
AI927207 AA782109 AW473233 AI804485 AW169216 AI572669 AA602182 AW015480 AW771865 AI270027 AA961816 AA283207 AI076962  
AI498487 AI348053 AI783914 H44405 AW799118 AA128330 AA515500 AA918281 W02156 AI905927 AA022701 W38382 R20795 T77861  
AW860878  
100528 45979\_1 BE386801 AU077299 AA143755 BE302747 AA853375 U30162 BE274163 BE277479 BE408180 BE274874 C15000 AA047476 N27099 AI359165  
AI638794 AI151283 AI863925 AW444977 AI207392 AA931263 AA443112 R40138 AW068538 AA351008 AA676972 R62503 AA916492 AW001865  
H42334 H38280 AA121497 AA114137 AI750938 M17783 AA383786 BE274462 AI753182 C05975 AA347404 AW069298 AI754351 AI754044  
AA188808 AA186879 AA565243 AL040655 AA456177 AI750722 AA045756 AA213580 C16936 AW578747 AW753731 H41632 N44761 R58560  
R61260 AA039902 N59721 AW992543 R68380 AA149686 T29017 H03739 BE383822 BE387105 BE408251 BE410425 H41560 AA247591  
BE389677 AI752233 AI566195 AA868004 AI424523 AW753720 AA852159 BE386803  
100559 2260\_1 NM\_000094 L02870 D13694 S51236 M96984 AW946290 M65158 AI285422 D29523 AL119886 AW630655 L06862 AI884355 AW168737 T29085  
AW797005 AW801340 AI355504 AW079048 AW801337 AI690455 AI972063 AW268565 W68588 AA587326 AA883498 AI033523 AW510356  
AW591998 H98463 AL043852 AI150055 AI566239 AI624803 AA844717 H40670 AA922334 AI864424 AW615094 AW451233 AI302203 F31221  
AI872170 W68589 AA904478 AI917631 AW014208 AW450759 AA847625 AI284033 AA848176 AA598507  
100576 9986\_1 X00356 NM\_001741 M26095 X03662 M12667 X02330 X02330 AA716058 AW296074 X04861 AI695720 AA719597  
124357 genbank\_N22401 N22401  
101624 entrez\_M55998M55998  
101625 entrez\_M57293M57293  
135158 57963\_1 AL037551 AI804716 AW439811 AI569470 AA075299 AI738572 AI270388 AI816783 AW263026 AI633951 AI655285 AI990572 AI950425  
AW241533 AA916883 AA576693 AA160156 AA613783 AW078884 AI888282 AI275241 AI133467 AA164921



Tables 2A-8C were previously filed on November 9, 2001 in USSN 60/339,245 (18501-004100US)

Table 2A shows 504 genes down-regulated in lung tumors relative to normal lung and chronically diseased lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigeneID: Unigene number  
 Unigene Title: Unigene gene title  
 R1: 90th percentile of AI for normal lung samples divided by the 80th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.  
 R2: median of AI for normal lung samples divided by 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.  
 R3: median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.  
 R4: average of AI for normal lung samples divided by average AI for squamous cell carcinoma and adenocarcinoma lung tumors.  
 R5: median of AI for normal lung samples divided by the 90th percentile of AI for adenocarcinomas.  
 R6: median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.  
 R7: average of AI for normal lung samples divided by the 90th percentile of AI for squamous cell carcinomas.  
 R8: median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for squamous cell carcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5	R6	R7	R8
100095	Z97171	Hs.78454	myocilin; trabecular meshwork inducible	40.20							
100115	NM_002084	Hs.336920	glutathione peroxidase 3 (plasma)								3.46
100138	U83508	Hs.2463	angiopoietin 1			2.30					
100299	D49493	Hs.2171	growth differentiation factor 10		11.00						
100306	U86749	Hs.80598	transcription elongation factor A (SII);						3.06		
100447	NM_014767	Hs.74583	KIAA0275 gene product								3.16
100458	S74019	Hs.247979	Vpre-B	42.40							
100862	AA005247	Hs.285754	Hepatocyte Growth Factor Receptor						4.13		
100959	AA359129	Hs.118127	actin; alpha; cardiac muscle				125.60				
101032	BE206854	Hs.46039	phosphoglycerate mutase 2 (muscle)	36.40							
101081	AF047347	Hs.4880	amyloid beta (A4) precursor protein-bind				34.60				
101088	X70697	Hs.553	solute carrier family 6 (neurotransmitter)				193.20				
101125	AJ250562	Hs.82749	transmembrane 4 superfamily member 2						3.10		
101180	U11874	Hs.846	interleukin 8 receptor; beta				54.86				
101308	L41390		"Homo sapiens core 2 beta-1,6-N-acetylgl	33.20							
101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do				36.40				
101345	NM_005795	Hs.152175	Calcitonin receptor-like			2.29					
101346	AI738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				70.55				
101397	M26380	Hs.180878	lipoprotein lipase								3.54
101414	NM_000066	Hs.38069	complement component 8; beta polypeptide							3.81	
101435	NM_001100	Hs.1288	actin; alpha 1; skeletal muscle				34.60				
101507	X16896	Hs.82112	interleukin 1 receptor; type I				37.60				
101530	M29874	Hs.1360	cytochrome P450; subfamily IIB (phenobar								4.25
101537	AI469059	Hs.184915	zinc finger protein; Y-linked			2.54					
101542	NM_000102	Hs.1363	cytochrome P450; subfamily XVII (steroid		5.50						
101545	BE246154	Hs.154210	EDG1; endothelial differentiation, sphin	39.40							
101554	BE207611	Hs.123078	thyroid stimulating hormone receptor		13.00						
101560	AW958272	Hs.83733	Intercellular adhesion molecule 2, exon								3.38
101574	M34182	Hs.158029	protein kinase; cAMP-dependent; catalyti						4.37		
101605	M37984	Hs.118845	troponin C; slow								3.80
101621	BE391804	Hs.62661	guanylate binding protein 1; interferon-	30.20							
101680	AA299330	Hs.1042	Sjogren syndrome antigen A1 (52kD; ribon							2.75	
101829	AW452398	Hs.129763	solute carrier family 8 (sodium/calcium						3.37		
101842	M93221	Hs.75182	mannose receptor; C type 1				38.20				
101961	AW004056	Hs.168357	"Hs-TBX2=T-box gene {T-box region} [huma			2.32					
101994	T92248	Hs.2240	uteroglobin								6.85
102020	AU077315	Hs.154970	transcription factor CP2			2.45					
102091	BE280901	Hs.83155	aldehyde dehydrogenase 7								6.75
102112	AW025430	Hs.155591	forkhead box F1	54.60							
102190	AA723157	Hs.73769	folate receptor 1 (adult)								3.98
102202	NM_000507	Hs.574	fructose-bisphosphatase 1								3.62
102241	NM_007351	Hs.268107	Multimerin			2.32					
102310	U33839		Accession not listed in Genbank		7.00						
102397	U41898		"Human sodium cotransporter RKST1 mRNA,	29.40							
102571	U60115	Hs.239069	"Homo sapiens skeletal muscle LIM-protei								3.75
102620	AA976427	Hs.121513	Human clone W2-6 mRNA from chromosome X						3.07		
102636	U67092		"Human ataxia-telangiectasia locus prote			2.40					
102667	U70867	Hs.83974	solute carrier family 21 (prostaglandin			3.15					
102675	U72512	Hs.7771	"Human B-cell receptor associated protei						3.56		
102698	M18667	Hs.1867	progastricsin (pepsinogen C)								4.51
102727	U79251	Hs.99902	opioid-binding protein/cell adhesion mol					12.00			
102852	V00571	Hs.75294	corticotropin releasing hormone	37.40							
103026	X54162	Hs.79386	thyroid and eye muscle autoantigen D1 (6					13.00			
103028	X54380	Hs.74094	pregnancy-zone protein	28.80							
103098	M86361		Human mRNA for T cell receptor; clone IG					10.00			
103117	X63578	Hs.295449	parvalbumin		6.00						
103241	X76223		H.sapiens MAL gene exon 4			2.47					
103280	U84722	Hs.76206	Cadherin 5, VE-cadherin (vascular epithe			2.69					
103360	Y16791	Hs.73082	keratin; hair; acidic; 5							2.16	



## WO 02/086443

## PCT/US02/12476

	103496	Y09267	Hs.132821	flavin containing monooxygenase 2					5.97
	103508	Y10141		"H.sapiens DAT1 gene, partial, VNTR"				3.27	
	103561	NM_001843	Hs.143434	contactin 1		2.40			
5	103569	NM_005512	Hs.151641	glycoprotein A repetitions predominant		2.99			
	103575	Z26256		"H.sapiens isoform 1 gene for L-type cal				4.18	
	103627	Z48513		H.sapiens XG mRNA (clone PEP6)				3.44	
	103767	BE244667	Hs.296155	CGI-100 protein					2.25
	103850	AA187101	Hs.213194	Hypothetical protein MGC10895; sim to SR			46.55		
10	104078	AA402801	Hs.303276	ESTs				3.05	
	104326	AW732858	Hs.143067	ESTs				3.54	
	104352	BE219898	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl				3.16	
	104398	AI423930	Hs.36790	ESTs; Weakly similar to putative p150 [H	64.80				
	104473	AI904823	Hs.31297	ESTs					3.38
15	104493	AW960427	Hs.79059	ESTs; Moderately similar to TGF-BETA REC		2.47			
	104495	AW975687	Hs.292979	ESTs	28.60				
	104595	AI799603	Hs.271568	ESTs				3.42	
	104597	AI364504	Hs.93967	ESTs; Weakly similar to Slit-1 protein [		6.00			
	104659	AW969769	Hs.105201	ESTs	34.00				
20	104686	AA010539	Hs.18912	ESTs		11.00			
	104691	U29690	Hs.37744	ESTs; Beta-1-adrenergic receptor	56.80				
	104764	AI039243	Hs.278585	ESTs			60.40		
	104776	AA026349		ESTs	34.20				
	104825	AA035613	Hs.141883	ESTs		3.03			
25	104865	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C	41.20				
	104942	NM_016348	Hs.10235	ESTs					3.27
	104989	R65998	Hs.285243	ESTs			40.00		
	105062	AW954355	Hs.36529	ESTs					3.20
	105101	H63202	Hs.38163	ESTs	34.20				
30	105173	U54617	Hs.8364	ESTs					4.17
	105194	R06780	Hs.19800	ESTs		16.00			
	105226	R58958	Hs.26608	ESTs					
	105256	AA430650	Hs.16529	transmembrane 4 superfamily member (tetr		2.34			
	105394	BE245812	Hs.8941	ESTs		2.72			
35	105647	Y09306	Hs.30148	homeodomain-interacting protein kinase 3	33.60				
	105789	AF106941	Hs.18142	arrestin; beta 2					3.59
	105817	AA397825		synaptopodin				4.46	
	105847	AW964490	Hs.32241	ESTs			35.40		
40	105894	AI904740	Hs.25691	calcitonin receptor-like receptor activi		3.43			
	105999	BE268786	Hs.21543	ESTs		7.00			
	106075	AA045290	Hs.25930	ESTs			42.60		
	106178	AL049935	Hs.301763	KIAA0554 protein	34.80				
	106381	AB040916	Hs.24106	ESTs			12.00		
45	106467	AA450040	Hs.154162	ADP-ribosylation factor-like 2				3.69	
	106536	AA329648	Hs.23804	ESTs			96.40		
	106569	R20909	Hs.300741	sorcin			47.20		
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr			220.40		
	106842	AF124251	Hs.26054	novel SH2-containing protein 3		2.55			
	106844	AA485055	Hs.158213	sperm associated antigen 6	39.20				
50	106870	AI983730	Hs.26530	serum deprivation response (phosphatidyl		2.28			
	106943	AW888222	Hs.9973	ESTs					4.28
	106954	AF128847	Hs.204038	ESTs					4.32
	107106	AA862496	Hs.28482	ESTs			10.45		
	107163	AF233588	Hs.27018	ESTs		2.57			
55	107201	D20378	Hs.30731	EST				3.84	
	107238	D59362	Hs.330777	EST		8.00			
	107376	U90545	Hs.327179	solute carrier family 17 (sodium phospho		10.67			
	107530	Y13622	Hs.85087	latent transforming growth factor beta b		2.32			
	107688	AW082221	Hs.60536	ESTs			34.60		
60	107706	AA015579	Hs.29276	ESTs	28.40				
	107723	AA015967		EST				3.29	
	107727	AA149707	Hs.173091	DKFZP434K151 protein			80.80		
	107750	AA017291	Hs.60781	ESTs			51.40		
	107751	AA017301	Hs.235390	ESTs				3.14	
65	107873	AK000520	Hs.143811	ESTs		9.00			
	107899	BE019261	Hs.83869	ESTs; Weakly similar to !!!! ALU SUBFAMI				3.65	
	107994	AA036811	Hs.48469	ESTs			44.60		
	107997	AL049176	Hs.82223	Human DNA sequence from clone 141H5 on c			32.00		
	108041	AW204712	Hs.61957	ESTs			30.80		
70	108048	AI797341	Hs.165195	ESTs					4.75
	108338	AA070773		"zm53g11.s1 Stratagene fibroblast (#9372		2.33			2.92
	108434	AA078899		"zm94b1.s1 Stratagene colon HT29 (#93722				3.06	
	108447	AA079126		"zm92a11.s1 Stratagene ovarian cancer (#					
	108480	AL133092	Hs.68055	ESTs			34.00		
75	108499	AA083103		"zn1b12.s1 Stratagene hNT neuron (#93723					3.36
	108535	R13949	Hs.226440	Homo sapiens clone 24881 mRNA sequence			19.00		
	108550	AA084867		"zn11f6.s1 Stratagene hNT neuron (#93723			12.00		
	108604	AA934589	Hs.49696	ESTs		2.33			
	108625	AW972330	Hs.283022	ESTs					5.82
80	108629	AA102425		"zn24c6.s1 Stratagene neuroepithelium NT				3.42	
	108655	AA099960		"zm65c6.s1 Stratagene fibroblast (#93721		7.00			
	108756	AA127221	Hs.117037	Homo sapiens mRNA; cDNA DKFZp564N1164 (f		6.05			
	108864	AI733852	Hs.199957	ESTs	28.80				
	108895	AL138272	Hs.62713	ESTs	32.80				
	108921	AI568801	Hs.71721	ESTs			57.80		
85	108967	AA142989	Hs.71730	ESTs	28.80				



## WO 02/086443

## PCT/US02/12476

5	109001	AI056548	Hs.72116	ESTs, Moderately similar to hedgehog-int	2.57				
	109003	AA147497	Hs.71825	ESTs				2.11	
	109004	AA156235	Hs.139077	EST	5.60				
	109065	AA161125	Hs.252739	EST			10.00		
	109250	H83784	Hs.62113	ESTs; Weakly similar to PHOSPHATIDYLETHA				3.44	
10	109490	AA233416	Hs.139202	ESTs				2.92	
	109510	AI798863	Hs.87191	ESTs	2.40				
	109578	F02208	Hs.27214	ESTs	10.00				
	109601	F02695	Hs.311662	EST			40.80		
	109613	H47315	Hs.27519	ESTs			54.40		
15	109650	R31770	Hs.23540	ESTs	31.20				
	109682	H18017	Hs.22869	ESTs	8.40				
	109724	D59899	Hs.127842	ESTs			29.40		
	109782	AB020644	Hs.14945	long fatty acyl-CoA synthetase 2 gene			8.00		
	109833	R79864	Hs.29889	ESTs	10.00				
20	109837	H00656	Hs.29792	ESTs	6.49				
	109977	T64183	Hs.282982	ESTs				2.75	
	109984	AI796320	Hs.10299	ESTs			107.00		
	110146	H41324	Hs.31581	ESTs; Moderately similar to SYNTAXIN 1B				2.22	
	110271	H28985	Hs.31330	ESTs				3.48	
25	110280	AW874263	Hs.32468	ESTs	44.20				
	110420	R93141	Hs.184261	ESTs			32.00		
	110578	T62507	Hs.11038	ESTs	28.40				
	110634	R98905	Hs.35992	ESTs			20.00		
	110726	AW961818	Hs.24379	potassium voltage-gated channel; shaker-				4.15	
30	110837	H03109	Hs.108920	ESTs; Weakly similar to semaphorin F [H.			56.80		
	110875	N35070	Hs.26401	tumor necrosis factor (ligand) superfam	3.13				
	110894	R92356	Hs.66881	ESTs; Moderately similar to cytoplasmic	5.33				
	110971	AI760098	Hs.21411	ESTs			44.60		
	111023	AV655386	Hs.7645	ESTs	32.40				
35	111057	T79639	Hs.14629	ESTs			17.14		
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f				4.58	
	111330	BE247767	Hs.18166	KIAA0870 protein				3.42	
	111374	BE250726	Hs.283724	ESTs; Moderately similar to HYA22 [H.sap				3.91	
	111442	AW449573	Hs.181003	ESTs			33.20		
40	111737	H04607	Hs.9218	ESTs			53.00		
	111747	AI741471	Hs.23666	ESTs	46.20				
	111807	R33508	Hs.18827	ESTs	16.00				
	111862	R37472	Hs.21559	EST				3.91	
	112045	AI372588	Hs.8022	TU3A protein				2.74	
45	112057	R43713	Hs.22945	EST				4.92	
	112214	AW148652	Hs.167398	ESTs			13.00		
	112263	R52393	Hs.25917	ESTs	2.43				
	112314	AW206093	Hs.748	ESTs	9.00				
	112324	R55965	Hs.26479	limbic system-associated membrane protei			14.00		
50	112362	AW300887	Hs.26638	ESTs; Weakly similar to CD20 receptor [H	2.49				
	112380	H63010	Hs.5740	ESTs	2.34				
	112425	AA324998	Hs.321677	ESTs; Weakly similar to !!!! ALU SUBFAMI	8.00				
	112473	R65993	Hs.279798	pregnancy specific beta-1-glycoprotein 9				4.53	
	112492	N51620	Hs.28694	ESTs			29.80	3.62	
55	112541	AF038392	Hs.116674	ESTs					
	112620	R80552	Hs.29040	ESTs	2.37				
	112623	AW373104	Hs.25094	ESTs	2.26				
	112867	T03254	Hs.167393	ESTs			12.00		
	112894	T08188	Hs.3770	ESTs	6.50				
60	112954	AA928953	Hs.6655	ESTs	7.00				
	113029	AW081710	Hs.7369	ESTs; Weakly similar to !!!! ALU SUBFAMI					4.39
	113086	AA346839	Hs.209100	DKFZP434C171 protein				4.47	
	113140	T50405	Hs.175967	ESTs			10.00		
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	14.00				
65	113257	AI821378	Hs.159367	ESTs				3.72	
	113394	T81473	Hs.177894	ESTs				3.60	
	113437	T85349	Hs.15923	EST	35.00				
	113454	AI022166	Hs.16188	ESTs	6.00				
	113502	T89130		ESTs	39.60				
70	113552	AI654223	Hs.16026	ESTs					3.88
	113645	T95358	Hs.333181	ESTs				2.58	
	113691	T96935	Hs.17932	EST			38.20		
	113706	AA004693	Hs.269192	ESTs				3.09	
	113883	U89281	Hs.11958	oxidative 3 alpha hydroxysteroid dehydro	2.31				
75	113924	BE178285	Hs.170056	Homo sapiens mRNA; cDNA DKFZp586B0220 (f	30.40				
	114035	W92798	Hs.269181	ESTs			13.00		
	114058	AK002016	Hs.114727	ESTs					5.00
	114084	AA708035	Hs.12248	ESTs			40.60		
	114121	H05785	Hs.25425	ESTs	2.31				
80	114124	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)	7.00				
	114275	AW515443	Hs.306117	interleukin 13 receptor; alpha 1	6.00				
	114297	AA149707	Hs.173091	DKFZP434K151 protein			48.80		
	114427	AA017176	Hs.33532	ESTs; Highly similar to Miz-1 protein [H				3.45	
	114449	AA020736		*ze63b11.s1 Soares retina N2b4HR Homo sa			10.00		
85	114452	AI369275	Hs.243010	ESTs, Moderately similar to RTC0_HUMAN G	14.00				
	114609	AA079505		*zm97a5.s1 Stratagene colon HT29 (#93722				3.13	
	114648	AA101056		*zn25b3.s1 Stratagene neuroepithelium NT			35.40		
	114731	BE094291	Hs.155651	Homo sapiens HNF-3beta mRNA for hepatocy					3.42
	114762	AA146979	Hs.288464	ESTs	33.00				



	114776	AA151719	Hs.95834	ESTs	34.40				
	115009	AA251561	Hs.48689	ESTs	30.20				
	115272	AW015947		ESTs; Weakly similar to hypothetical L1	32.60				
5	115279	AW964897	Hs.290825	ESTs		6.00		12.00	
	115302	AL109719	Hs.47578	ESTs					3.32
	115365	AW976252	Hs.268391	ESTs					
	115559	AL079707	Hs.207443	ESTs			48.00		
	115566	AI142336	Hs.43977	ESTs			56.20		
10	115683	AF255910	Hs.54650	ESTs; Weakly similar to (define not ava	31.40				
	115744	AA418538	Hs.43945	ESTs; Highly similar to dJ1178H5.3 [H.sa				33.60	
	115819	AA486620	Hs.41135	Endomucin 2				74.40	
	115949	AI478427	Hs.43125	ESTs		3.18			
	115965	AA001732	Hs.173233	ESTs			388.80		
15	116035	AA621405	Hs.184664	ESTs			33.20		
	116049	AA454033	Hs.41644	ESTs			45.80		
	116081	AI190071	Hs.55278	ESTs					3.57
	116082	AB029496	Hs.59729	ESTs		3.06			
	116213	AA292105	Hs.326740	leucine rich repeat (in FLII) interactin	50.60				
20	116228	AI767947	Hs.50841	ESTs; Weakly similar to tuftelin [M.musc		3.85			
	116250	N76712	Hs.44829	ESTs		6.00			
	116419	AI613480	Hs.47152	ESTs; Weakly similar to testicular tekti			30.00		
	116617	D80761	Hs.45220	EST		2.27			
	116784	AB007979	Hs.301281	tenascin R (restrictin; janusin)	47.20				
25	116835	N39230	Hs.38218	ESTs			41.20		
	116970	AB023179	Hs.9059	KIAA0962 protein				11.00	
	117023	AW070211	Hs.102415	ESTs			91.00		
	117027	AW085208	Hs.130093	ESTs	49.40				
	117036	H88908	Hs.41192	EST			32.60		
30	117110	AA160079	Hs.172932	ESTs		8.67			
	117209	W03011	Hs.306881	ESTs			30.60		
	117325	N23599	Hs.43396	ESTs				9.29	
	117454	N29569	Hs.44055	ESTs					3.19
	117475	N30205	Hs.93740	ESTs	44.00				
35	117543	BE219453	Hs.42722	ESTs		16.00			
	117567	AW444761	Hs.44565	ESTs				12.00	
	117570	N48649	Hs.44583	ESTs				11.00	
	117600	N34963	Hs.44676	EST					3.74
	117730	N45513	Hs.46608	ESTs		6.00			
40	117791	N48325	Hs.93956	EST		9.00			
	117929	N51075	Hs.47191	ESTs			29.20		
	117990	AA446167	Hs.47385	ESTs		8.00			
	118224	N62275	Hs.48503	EST	31.40				
	118244	N62516	Hs.48556	ESTs	32.80				
45	118357	AL109667	Hs.124154	Homo sapiens mRNA full length insert cDN		2.40			
	118446	N66361	Hs.269121	ESTs		2.28			
	118447	N66399	Hs.49193	EST	30.80				
	118530	N67900	Hs.118446	ESTs					3.10
	118549	N68163	Hs.322954	EST					3.41
50	118823	W03754	Hs.50813	ESTs; Weakly similar to long chain fatty		3.94			
	118862	W17065	Hs.54522	ESTs					3.58
	118935	AI979247	Hs.247043	KIAA0525 protein			33.00		
	118944	AI734233	Hs.226142	ESTs; Weakly similar to !!!! ALU SUBFAMI				11.43	
	118995	N94591	Hs.323056	ESTs		14.00			
55	119073	BE245360	Hs.279477	ERG-2/ERG-1; V-ets avian erythroblastosi			52.60		
	119268	T16335	Hs.65325	EST	31.40				
	119514	W37937		Accession not listed in Genbank					3.50
	119824	W74536	Hs.184	advanced glycosylation end product-speci		2.75			
	119831	AL117664	Hs.58419	DKFZP586L2024 protein					3.21
60	119861	W78816	Hs.49943	ESTs; Moderately similar to !!!! ALU SUB			33.80		
	119889	W84346	Hs.58671	ESTs			30.03		
	119921	W86192	Hs.58815	ESTs	29.00				
	120082	H80286	Hs.40111	ESTs					3.80
	120094	AA811339	Hs.124049	ESTs		6.00			
65	120132	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)			36.60		
	120378	AA223249	Hs.285728	ESTs		12.00			
	120404	AB023230	Hs.96427	KIAA1013 protein	39.40				
	120504	AA256837		ESTs				8.00	
	120512	N55761	Hs.194718	ESTs	33.00				
70	120667	AA287740	Hs.78335	microtubule-associated protein; RP/EB fa					4.18
	120777	AA287702	Hs.10031	KIAA0955 protein			46.60		
	121082	AA398722		ESTs			39.00		
	121191	AA400205	Hs.104447	ESTs	41.60				
	121248	AA400914	Hs.97827	EST					5.08
75	121363	AI287280	Hs.97933	ESTs				12.00	
	121366	AI743515		ESTs				20.00	
	121483	AI660332	Hs.25274	ESTs; Moderately similar to putative sev					3.32
	121518	AA412155		ESTs			30.20		
	121545	AA412442	Hs.98132	ESTs		2.29			
80	121622	AA416931	Hs.126065	ESTs		9.00			
	121665	AA416556	Hs.98234	ESTs			34.80		
	121709	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	34.80				
	121730	AI140683	Hs.98328	ESTs	38.80				
	121740	AA421138	Hs.98334	EST		7.00			
85	121772	AI590770	Hs.110347	Homo sapiens mRNA for alpha integrin bin	36.20				
	121821	AL040235	Hs.3346	ESTs					3.61



**PCT/US02/12476**

99



	129210	AL039940	Hs.202949	KIAA1102 protein					4.09
	129240	AA361258	Hs.237868	interleukin 7 receptor	2.29				
	129262	BE222198	Hs.109843	ESTs			3.30		
5	129301	AF182277	Hs.330780	Human cytochrome P450-IIB (hIIB3) mRNA;					4.05
	129331	AW167668	Hs.279772	ESTs; Highly similar to CGI-38 protein [					4.09
	129381	AW245805	Hs.110903	claudin 5 (transmembrane protein deleted	2.93				
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		160.80			
	129595	U09550	Hs.1154	oviductal glycoprotein 1; 120kD			10.00		
10	129613	AW978517	Hs.172847	ESTs; Weakly similar to collagen alpha 1				3.40	
	129782	AW016932	Hs.104105	EST	9.00				
	129950	F07783	Hs.1369	decay accelerating factor for complement		87.80			
	129958	R27496	Hs.1378	annexin A3		44.60			
	129959	AL036554	Hs.274463	defensin; alpha 1; myeloid-related seque	2.72				
15	130160	AA305688	Hs.267695	UDP-Gal:betaGlcNAc beta 1;3-galactosyltr		42.20			
	130259	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator	2.54				
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic; Dr		51.60			
	130312	AF056195	Hs.15430	DKFZP586G1219 protein				3.16	
	130436	NM_001928	Hs.155597	D component of complement (adipsin)					4.11
20	130523	AA999702	Hs.214507	ESTs				4.77	
	130799	AB028945	Hs.12696	ESTs	6.00				
	130885	NM_005883	Hs.20912	adenomatous polyposis coli like				3.54	
	131002	AL050295	Hs.22039	KIAA0758 protein					3.50
	131012	AL039940	Hs.202949	KIAA1102 protein	20.00				
25	131031	NM_001650	Hs.288650	aquaporin 4	41.20				
	131061	N64328	Hs.268744	ESTs; Moderately similar to KIAA0273 [H.		31.40			
	131066	AW169287	Hs.22588	ESTs		29.60			
	131082	AI091121	Hs.246218	ESTs; Weakly similar to zinc finger prot			9.00		
	131087	AF147709	Hs.22824	ESTs; Weakly similar to p160 myb-binding					3.86
30	131161	AF033382	Hs.23735	potassium voltage-gated channel; subfami				3.14	
	131179	AA171388	Hs.184482	DKFZP586D0624 protein				3.80	
	131182	AI824144	Hs.23912	ESTs					3.67
	131205	NM_003102	Hs.2420	superoxide dismutase 3; extracellular	2.98				
	131277	AA131466	Hs.23767	ESTs	3.15				
35	131281	AA251716	Hs.25227	ESTs		32.20			
	131282	X03350	Hs.4	alcohol dehydrogenase 3 (class I); gamma					3.44
	131285	AI567943	Hs.25274	ESTs; Moderately similar to putative sev				6.40	
	131355	R52804	Hs.25956	DKFZP564D206 protein	8.00				
	131391	AW085781	Hs.26270	ESTs	10.00				
40	131461	AA992841	Hs.27263	butyrate response factor 2 (EGF-response	28.80				
	131487	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f					4.03
	131517	AB037789	Hs.263395	ESTs; Highly similar to semaphorin VIa [	39.00				
	131545	AL137432	Hs.28564	ESTs			11.00		
	131583	AK000383	Hs.323092	ESTs; Weakly similar to dual specificity			10.00		
45	131647	AA359615	Hs.30089	ESTs	2.47				
	131675	H15205	Hs.30509	ESTs				3.06	
	131676	AI126821	Hs.30514	ESTs	45.80				
	131708	S60415	Hs.30941	calcium channel; voltage-dependent; beta	2.28				
	131717	X94630	Hs.3107	CD97 antigen					3.78
50	131756	AA443966	Hs.31595	ESTs		40.60			
	131762	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII inhi					3.67
	131821	AA017247	Hs.164577	ESTs	2.87				
	131839	AB014533	Hs.33010	KIAA0633 protein					3.48
	131861	AL096858	Hs.184245	KIAA0929 protein Msx2 interacting nuclea	54.00				
55	132015	AI418006	Hs.3731	ESTs		49.20			
	132070	BE622641	Hs.38489	ESTs		34.80			
	132242	AA332697	Hs.42721	ESTs	2.68				
	132334	AW080704	Hs.45033	lacrima proline rich protein	4.66				
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	34.20				
60	132490	NM_001290	Hs.4980	LIM binding domain 2	2.66				
	132533	AI922988	Hs.172510	ESTs	13.00				
	132598	X80031	Hs.530	collagen; type IV; alpha 3 (Goodpasture		30.60			
	132619	H28855	Hs.53447	ESTs; Moderately similar to kinesin ligh				4.02	
	132652	N41739	Hs.61260	ESTs				3.18	
65	132726	N52298	Hs.55608	ESTs; Weakly similar to cDNA EST yk484g1			11.43		
	133028	R51604	Hs.300842	ESTs	2.37				
	133071	BE384932	Hs.64313	ESTs	2.27				
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein	2.63				
	133129	AA428580	Hs.65551	ESTs					5.49
70	133147	AA026533	Hs.66	interleukin 1 receptor-like 1	6.20				
	133151	NM_014051	Hs.94896	ESTs				3.69	
	133213	AA903424	Hs.6786	ESTs		31.40			
	133276	AW978439	Hs.69504	ESTs			9.00		
	133377	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene famil	41.20				
75	133407	AF017987	Hs.7306	secreted frizzled-related protein 1	50.20				
	133535	AL134030	Hs.284180	protocadherin 2 (cadherin-like 2)				3.72	
	133537	U41518	Hs.74602	aquaporin 1 (channel-forming integral pr					3.35
	133656	BE149455	Hs.75415	Accession not listed in Genbank	2.65				
	133689	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)		90.80			
80	133779	T58486	Hs.222566	ESTs				3.05	
	133978	AF035718	Hs.78061	transcription factor 21	2.92				
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec					3.45
	134000	AW175787	Hs.334841	selenium binding protein 1					4.05
	134111	AI372588	Hs.8022	TU3A protein	4.49				
85	134185	AA285136	Hs.301914	Homo sapiens mRNA; cDNA DKFZp586K1220 (f				3.27	
	134204	AI873257	Hs.7994	ESTs; Weakly similar to CGI-69 protein [		40.80			



**WO 02/086443**

**PCT/US02/12476**

	134641	AI092634	Hs.156114	protein tyrosine phosphatase; non-recept				3.76
	134677	AA251363	Hs.177711	ESTs		32.20		
	134745	NM_000685	Hs.89472	angiotensin receptor 1B	15.00			
5	134749	T28499	Hs.89485	carbonic anhydrase IV		3.05		
	134786	T29618	Hs.89640	angiopoietin 1 receptor; TEK tyrosine ki			57.80	
	134825	U33749	Hs.197764	thyroid transcription factor 1				3.73
	134978	AI829008	Hs.333383	ficolin (collagen/fibrinogen domain-cont		2.52		
	135010	N50465	Hs.92927	ESTs			31.60	
	135053	AW796190	Hs.93678	ESTs				3.21
10	135081	AF069517	Hs.173993	RNA binding motif protein 6	28.80			
	135091	AA493650	Hs.94367	ESTs				4.24
	135135	AA775910	Hs.95011	syntrophin; beta 1 (dystrophin-associate		8.00		
	135203	C15737	Hs.269386	ESTs				4.31
	135236	AI636208	Hs.96901	ESTs	43.00			
15	135266	R41179	Hs.97393	Human mRNA for KIAA0328 gene; partial cd				6.42
	135346	NM_000928	Hs.992	phospholipase A2; group IB (pancreas)		3.82		
	135378	AW961818	Hs.24379	potassium voltage-gated channel; shaker-		4.15		
	135387	NM_001972	Hs.99863	elastase 2; neutrophil	37.20			
	135388	W27965	Hs.99865	EST	38.80			
20	135402	L12398	Hs.99922	dopamine receptor D4				4.21

25 TABLE 2B shows the accession numbers for those primekeys lacking unigenelD's for Table 2A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

30 Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
35	108447	43452_-7 AA079126
	108550	120073_1 AA084867 AA084996
	108655	127522_1 AA099960 AA113013
	102397	44371_-1 U41898
	126303	1525933_1 D78841 D78880
40	125810	1554054_1 H00083 R81062
	103627	2615_2 Z48513 Z48512
	121366	280401_1 AI743515 AA405617 AW276706
	114609	116777_1 AA079505 AA079537
	115272	172113_1 AWO15947 AA211890 AA279425
	108338	112186_1 AA070773 AA070774
45	108434	114012_1 AA078899 AA078782 AA075788
	123802	genbank_AA620448 AA620448
	102310	NOT_FOUND_entrez_U33839 U33839
	102636	entrez_U67092 U67092
	104776	genbank_AA026349 AA026349
50	120504	genbank_AA256837 AA256837
	113502	genbank_T89130T89130
	108499	genbank_AA083103 AA083103
	101308	entrez_L41390 L41390
	108629	genbank_AA102425 AA102425
55	103098	221_215 M86361 Z26593 X02850 D13070 AE000659 M17649 M87869 M87871 X61077 M16286 AF018169 X61079 S59351 X60142 AF043169
	103241	entrez_X76223 X76223
	103508	entrez_Y10141 Y10141
	103575	entrez_Z26256 Z26256
	119514	NOT_FOUND_entrez_W37937 W37937
60	121082	genbank_AA398722 AA398722
	128634	AA464918_at AA464918
	105817	genbank_AA397825 AA397825
	121518	genbank_AA412155 AA412155
	114449	genbank_AA020736 AA020736
65	114648	genbank_AA101056 AA101056
	121950	genbank_AA429515 AA429515
	107723	genbank_AA015967 AA015967



Table 3A shows 452 genes up-regulated in chronically diseased lung relative to normal lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number					
	ExAccn:	Exemplar Accession number, Genbank accession number					
	UnigenelD:	Unigene number					
	Unigene Title:	Unigene gene title					
10	R1:	80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of AI for normal lung samples.					
	R2:	80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas					
	R3:	70th percentile of AI for chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples					
15							
	Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2	R3
20	135423	U50531	Hs.138751	Human BRCA2 region, mRNA sequence CG030	12.40		
	135378	AW961818	Hs.24379	MUM2 protein			2.13
	135346	NM_000928	Hs.992	phospholipase A2, group 1B (pancreas)			
	135235	AW298244	Hs.293507	ESTs	12.40		
	135057	U90268	Hs.93810	cerebral cavernous malformations 1	11.67		
25	134951	BE305081	Hs.169358	hypothetical protein		8.00	
	134799	M36821	Hs.89690	GRO3 oncogene		8.20	
	134786	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous			
	134772	NM_000829	Hs.163697	glutamate receptor, ionotropic, AMPA 4	29.80		
	134752	BE246762	Hs.89499	arachidonate 5-lipoxygenase			1.93
	134749	T28499	Hs.89485	carbonic anhydrase IV			2.07
30	134696	BE326276	Hs.8861	ESTs			
	134636	NM_005582	Hs.87205	lymphocyte antigen 64 (mouse) homolog, r	13.60		
	134627	AI018768	Hs.12482	glyceronephosphate O-acyltransferase			1.92
	134622	AW975159	Hs.293097	ESTs, Weakly similar to A55380 faciogeni			1.92
35	134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	13.20		
	134561	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h			1.78
	134468	NM_001772	Hs.83731	CD33 antigen (gp67)		6.20	
	134417	NM_006416	Hs.82921	solute carrier family 35 (CMP-sialic aci			
	134343	D50683	Hs.82028	transforming growth factor, beta recepto			
40	134323	BE170651	Hs.8700	deleted in liver cancer 1			
	134300	NM_001430	Hs.8136	endothelial PAS domain protein 1			
	134299	AW580939	Hs.97199	complement component C1q receptor			
	134253	X52075	Hs.80738	sialophorin (gpL115, leukosialin, CD43)	20.60		
	134182	D52059	Hs.7972	KIAA0871 protein	12.20		
45	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec			
	133978	AF035718	Hs.78061	transcription factor 21			
	133835	AI677897	Hs.76640	RGC32 protein			
	133651	AI301740	Hs.173381	dihydropyrimidinase-like 2			
	133633	D21262	Hs.75337	nucleolar and coiled-body phosphprotein	15.20		
50	133565	AW955776	Hs.313500	ESTs, Moderately similar to ALU7_HUMAN A			
	133548	AW946384	Hs.178112	DNA segment, single copy probe LNS-CAI/L			1.77
	133488	AA335295	Hs.74120	adipose specific 2			
	133478	X83703	Hs.31432	cardiac ankyrin repeat protein			2.08
	133337	AF085983	Hs.293676	ESTs		9.60	
55	133200	AB037715	Hs.183639	hypothetical protein FLJ10210			1.77
	133153	AF070592	Hs.66170	HSKM-B protein	30.60		
	133130	AI128606	Hs.6557	zinc finger protein 161	22.60		
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein			
	132928	AW168082	Hs.169449	protein kinase C, alpha	13.80		
60	132836	AB023177	Hs.29900	KIAA0960 protein			
	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,	41.60		
	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypotheti	40.40		
	132548	X12830	Hs.193400	interleukin 6 receptor		7.20	
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi		4.76	
65	132439	AK001942	Hs.4863	hypothetical protein DKFZp566A1524			1.88
	132240	AB018324	Hs.42676	KIAA0781 protein	21.20		
	132210	NM_007203	Hs.42322	A kinase (PRKA) anchor protein 2			1.99
	132199	AL041299	Hs.165084	ESTs	15.20		
	131751	T96555	Hs.31562	ESTs			1.76
70	131745	AI828559	Hs.31447	ESTs, Moderately similar to A46010 X-li	27.80		
	131694	NM_000246	Hs.3076	MHC class II transactivator		4.00	
	131686	NM_012296	Hs.30687	GRB2-associated binding protein 2			
	131676	AI126821	Hs.30514	ESTs		6.20	
	131629	Z45794	Hs.238809	ESTs	21.40		
75	131589	C18825	Hs.29191	epithelial membrane protein 2			
	131536	AA019201	Hs.269210	ESTs		9.40	
	131517	AB037789	Hs.263395	sema domain, transmembrane domain (TM),		3.59	
	131355	R52804	Hs.25956	DKFZP564D206 protein		4.48	
	131253	R71802	Hs.24853	ESTs	15.00		
80	131207	AF104266	Hs.24212	latrophilin			1.75
	131156	AI472209	Hs.323117	ESTs			1.84
	131066	AW169287	Hs.22588	ESTs		3.54	
	131061	N64328	Hs.268744	KIAA1796 protein			
	131053	AA348541	Hs.296261	guanine nucleotide binding protein (G pr			1.93
85	130895	AA641767	Hs.21015	hypothetical protein DKFZp564L0864 simil	16.60		
	130762	D84371	Hs.1898	paraoxonase 1	12.00		



	130657	AW337575	Hs.201591	ESTs		
	130655	AI831962	Hs.17409	cysteine-rich protein 1 (intestinal)		
	130589	AL110226	Hs.16441	DKFZP434H204 protein		2.08
5	130562	D50402	Hs.182611	solute carrier family 11 (proton-coupled		1.91
	130555	R69743	Hs.116774	integrin, alpha 1	9.60	
	130365	W56119	Hs.155103	eukaryotic translation initiation factor	11.60	
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic, Dr	6.60	
	130259	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator		1.91
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 18)	21.20	
	129958	R27496	Hs.1378	annexin A3	5.05	
	129898	AI672731	Hs.13256	ESTs		
	129875	AA181018	Hs.13056	hypothetical protein FLJ13920	18.60	
	129699	AB007899	Hs.12017	homolog of yeast ubiquitin-protein ligas		
15	129626	F13272	Hs.111334	ferritin, light polypeptide		
	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	22.63	
	129593	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f		
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		2.53
	129527	AA769221	Hs.270847	delta-tubulin	39.20	
20	129402	W72062	Hs.11112	ESTs		2.11
	129385	AA172106	Hs.110950	Rag C protein	15.20	
	129315	NM_014563	Hs.174038	spondyloepiphyseal dysplasia, late	12.40	
	129312	T97579	Hs.110334	ESTs, Weakly similar to I78885 serine/th	20.83	
	129240	AA361258	Hs.237868	interleukin 7 receptor		1.95
25	129210	AL039940	Hs.202949	KIAA1102 protein		
	129122	AW958473	Hs.301957	nudix (nucleoside diphosphate linked moi	4.20	
	129057	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)		
	128946	Y13153	Hs.107318	kynurenine 3-monooxygenase (kynurenine 3	5.20	
	128798	AF015525	Hs.302043	chemokine (C-C motif) receptor-like 2		
30	128789	AW368576	Hs.139851	caveolin 2		2.24
	128778	AA504776	Hs.186709	ESTs, Weakly similar to I38022 hypothet	12.20	
	128766	AW160432	Hs.296460	craniofacial development protein 1	26.40	
	128631	R44238	Hs.155546	KIAA1080 protein; Golgi-associated, gamm		1.78
	128624	BE154765	Hs.102647	ESTs, Weakly similar to TRHY_HUMAN TRICH		2.51
35	128609	NM_003616	Hs.102456	survival of motor neuron protein interac	16.00	
	128603	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	12.80	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel, s		4.00
	128458	H55864	Hs.56340	ESTs		
40	128061	AF150882	Hs.186877	sodium channel, voltage-gated, type XII,	17.20	
	127968	AA830201	Hs.124347	ESTs	21.30	
	127959	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L		
	127944	AI557081	Hs.262476	S-adenosylmethionine decarboxylase 1	10.60	
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	13.40	
45	127896	AI669586	Hs.222194	ESTs		7.00
	127859	AA761802	Hs.291559	ESTs	14.00	
	127817	AA836641	Hs.163085	ESTs	14.00	
	127742	AW293496	Hs.180138	ESTs	11.00	
	127628	AI240102	Hs.322430	NDRG family, member 4	11.10	
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture		
50	127582	AA908954	Hs.130844	ESTs	19.60	
	127543	AK000787	Hs.157392	Homo sapiens cDNA FLJ20780 fis, clone CO	15.40	
	127535	AA568424	Hs.164450	ESTs	17.50	
	127404	AI379920	Hs.270224	ESTs	14.60	
	127396	L31968	Hs.187991	DKFZP564A122 protein	15.40	
55	127374	AA442797	Hs.312110	ESTs, Weakly similar to I38022 hypothet	14.60	
	127346	AA203616	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe	21.00	
	127340	BE047653	Hs.119183	ESTs, Weakly similar to ZN91_HUMAN ZINC	15.80	
	127307	AW962712	Hs.126712	ESTs, Weakly similar to AF191020 1 E2IG5		
	127242	AW390395	Hs.181301	cathepsin S	22.60	
60	127167	AA625690	Hs.190272	ESTs	21.40	
	127046	AA321948	Hs.293968	ESTs	41.20	
	126928	AA480902	Hs.137401	ESTs	11.00	
	126900	AF137386	Hs.12701	plasmolipin		1.78
65	126852	AA399961		gb:zu68c01.r1 Soares_testis_NHT Homo sap	5.60	
	126816	AA248234		gb:csg2228.seq.F Human fetal heart, Lamb	12.20	
	126812	AB037860	Hs.173933	nuclear factor I/A	17.19	
	126666	AA648886	Hs.151999	ESTs	13.57	
	126645	AA316181	Hs.61635	six transmembrane epithelial antigen of	15.40	
70	126592	AI611153	Hs.6093	Homo sapiens cDNA: FLJ22783 fis, clone K		4.67
	126556	AF255303	Hs.112227	membrane-associated nucleic acid binding	18.00	
	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	16.77	
	126299	AW979155	Hs.298275	amino acid transporter 2	14.60	
	126218	AL049801	Hs.13649	Novel human gene mapping to chromosome 13	3.50	
75	126182	AA721331	Hs.293771	ESTs	13.40	
	126177	AW752782	Hs.129750	hypothetical protein FLJ10546	18.20	
	126142	H86261	Hs.40568	ESTs	14.00	
	126077	M78772	Hs.210836	ESTs	16.59	
	125994	AI990529	Hs.270799	ESTs	17.40	
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	13.00	
80	125847	AW161885	Hs.249034	ESTs	49.57	
	125831	H04043		gb:yj45c03.r1 Soares placenta Nb2HP Homo		
	125731	R61771	Hs.26912	ESTs	13.20	
	125676	BE612918	Hs.151973	hypothetical protein FLJ23511	11.20	
	125561	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S		
85	125552	H09701	Hs.278366	ESTs, Weakly similar to I38022 hypotheti	12.60	
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A	33.40	



	125422	AA903229	Hs.153717	ESTs			1.80
	125331	AI422996	Hs.161378	ESTs	38.00		
	125309	T12411	Hs.183745	hypothetical protein FLJ13456	18.20		
5	125167	AL137540	Hs.102541	netrin 4			1.95
	125139	AW194933	Hs.9788	hypothetical protein MGC10924 similar to			1.84
	125042	T78906	Hs.269432	ESTs, Moderately similar to ALU1_HUMAN	21.80		
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl		10.60	
	124631	NM_014053	Hs.270594	FLVCR protein	23.20		
	124578	N68321	Hs.231500	EST	21.43		
10	124574	AL036596	Hs.42322	A kinase (PRKA) anchor protein 2			1.77
	124472	N52517	Hs.102670	EST	37.20		
	124438	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A			
	124357	N22401		gb:yw37g07.s1 Morton Fetal Cochlea Homo	14.64		
15	124306	AW973078	Hs.293039	ESTs		4.00	
	124214	H58608	Hs.151323	ESTs			
	124097	AW298235	Hs.101689	ESTs		27.20	
	123978	T89832	Hs.170278	ESTs			2.03
	123972	T46848	Hs.70337	immunoglobulin superfamily, member 4		6.00	
20	123961	AL050184	Hs.21610	DKFZP434B203 protein			1.79
	123936	NM_004673	Hs.241519	angiopoietin-like 1		15.80	
	123802	AA620448		gb:ae58c09.s1 Stratagene lung carcinoma		4.23	
	123734	AA609861	Hs.312447	ESTs		4.20	
	123619	AA602964		gb:no97c02.s1 NCI_CGAP_Pr2 Homo sapiens	33.60		
25	123596	AA421130	Hs.112640	EST	10.93		
	123476	AA384564	Hs.108829	ESTs			2.18
	123340	AA504264	Hs.182937	peptidylprolyl isomerase A (cyclophilin	11.20		
	123190	AA489212	Hs.105228	EST	14.20		
	123136	AW451999	Hs.194024	ESTs		7.00	
30	123073	AA485061	Hs.105652	ESTs	31.20		
	123055	AA482005	Hs.105102	ESTs, Weakly similar to reverse transcri		4.80	
	122699	AA456130	Hs.301721	KIAA1255 protein		5.00	
	122679	AA811286	Hs.192837	ESTs, Weakly similar to ALU5_HUMAN ALU S	14.40		
	122633	NM_001546	Hs.34853	inhibitor of DNA binding 4, dominant neg			
35	122553	AA451884	Hs.190121	ESTs	40.00		
	122544	AW973253	Hs.292689	ESTs	15.40		
	122485	AA524547	Hs.160318	FXRD domain-containing ion transport reg			1.81
	122211	AA300900	Hs.98849	ESTs, Moderately similar to AF161511 1 H		12.10	
	122127	AW207175	Hs.106771	ESTs			1.95
40	122011	AA431082		gb:zw78a10.s1 Soares_testis_NHT Homo sap			1.89
	121992	AI860775	Hs.98506	ESTs		3.60	
	121989	W56487	Hs.193784	Homo sapiens mRNA; cDNA DKFZp586K1922 (f			2.01
	121835	AB033030	Hs.300670	KIAA1204 protein			1.85
	121726	AF241254	Hs.178098	angiotensin I converting enzyme (peptidy	12.43		
45	121690	AV660305	Hs.110286	ESTs			1.82
	121643	AA640987	Hs.193767	ESTs			
	121633	AA417011	Hs.98175	EST	14.00		
	121622	AA416931	Hs.126065	ESTs		16.40	
	121497	AA412031	Hs.97901	EST	11.20		
50	121351	AW206227	Hs.287727	hypothetical protein FLJ23132	12.20		
	121314	W07343	Hs.182538	phospholipid scramblase 4			1.83
	121242	AA400857	Hs.97509	ESTs	22.40		
	121059	AA393283		gb:zt74e03.r1 Soares_testis_NHT Homo sap	14.80		
	120934	AA226198		gb:nc26a07.s1 NCI_CGAP_Pr1 Homo sapiens	21.20		
55	120755	AA312934	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone			1.79
	120637	AA811804		gb:ob39a05.s1 NCI_CGAP_GCB1 Homo sapiens	20.00		
	120484	AA253170	Hs.96473	EST	40.20		
	120336	N85785	Hs.181165	eukaryotic translation elongation factor		6.60	
	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	16.80		
60	120132	W57554	Hs.125019	ESTs		4.73	
	120041	AA830882	Hs.59368	ESTs			1.75
	119996	W88996	Hs.59134	EST		7.20	
	119970	AA767718	Hs.93581	hypothetical protein FLJ10512	11.20		
	119861	W78816	Hs.49943	ESTs, Weakly similar to S65657 alpha-1C-		3.78	
65	119824	W74536	Hs.184	advanced glycosylation end product-speci			
	119740	AW021407	Hs.21068	hypothetical protein	20.20		
	119271	AI061118	Hs.65328	Fanconi anemia, complementation group F	15.20		
	119221	C14322	Hs.250700	tryptase beta 1			
	119126	R45175	Hs.117183	ESTs	12.60		
70	119073	BE245360	Hs.279477	ESTs			
	118928	AA312799	Hs.283689	activator of CREM in testis		10.00	
	118901	AW292577	Hs.94445	ESTs		3.96	
	118661	AL137554	Hs.49927	protein kinase NYD-SP15		9.60	
	118607	AI377444	Hs.54245	ESTs, Weakly similar to S65824 reverse t	10.40		
75	118449	AI813865	Hs.164478	hypothetical protein FLJ21939 similar to			1.90
	118416	N66028	Hs.49105	FKBP-associated protein	16.20		
	118379	N64491	Hs.48990	ESTs		4.00	
	118329	N63520		gb:yy62f01.s1 Soares_multiple_sclerosis_		6.60	
	118320	N63451	Hs.141600	ESTs, Weakly similar to alternatively s		3.80	
80	118253	AA497044	Hs.20887	hypothetical protein FLJ10392	17.60		
	118124	N56968	Hs.46707	chromosome 21 open reading frame 37	14.00		
	118056	AB037746	Hs.42768	hypothetical protein DKFZp761O0113			1.86
	118032	N52802	Hs.47544	EST		5.00	
	117840	T26379	Hs.48802	Homo sapiens clone 23632 mRNA sequence		4.00	
85	117404	N39725	Hs.15220	zinc finger protein 106			1.90
	117314	N32498	Hs.42829	ESTs	14.20		



	117209	W03011	Hs.306881	MSTP043 protein			
	117023	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f		2.31	
	116814	H50834		gb:yp86a10.s1 Soares fetal liver spleen	20.20		
5	116784	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific		3.51	
	116766	AI608657	Hs.95097	ESTs	16.20		
	116712	AW901618	Hs.61935	Homo sapiens mRNA; cDNA DKFZp761I071 (fr		6.80	
	116707	H10344	Hs.49050	ESTs, Weakly similar to A Chain A, Human	18.60		
	116351	AL133623	Hs.82501	similar to mouse Xrn1 / Dhms2 protein	19.40		
10	116279	AW971248	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S			
	116166	AL039940	Hs.202949	KIAA1102 protein		2.13	
	116152	AL040521	Hs.15220	zinc finger protein 106		1.75	
	116117	BE613410	Hs.31575	SEC63, endoplasmic reticulum translocon	13.20		
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	30.11		
15	115965	AA001732	Hs.173233	hypothetical protein FLJ10970		2.36	
	115955	AF263613	Hs.44198	intracellular membrane-associated calciu	18.20		
	115844	AI373062	Hs.332938	hypothetical protein MGC5370	18.57		
	115683	AF255910	Hs.54650	junctional adhesion molecule 2		23.00	
	115673	AA406341	Hs.269908	Homo sapiens cDNA FLJ11991 fis, clone HE	11.82		
20	115672	AI889110	Hs.73251	ESTs	10.60		
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1		1.76	
	115313	AA808001	Hs.184411	albumin	25.20		
	115279	AW964897	Hs.290825	ESTs		8.00	
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L		1.80	
25	115110	AK001671	Hs.11387	KIAA1453 protein	14.20		
	114999	BE246481	Hs.87856	ESTs	19.20		
	114930	AA237022	Hs.188717	ESTs		5.60	
	114922	AA235672	Hs.87491	ESTs		3.60	
	114837	BE244930	Hs.166895	ESTs	43.70		
30	114769	AA149060	Hs.296100	ESTs	11.00		
	114761	AA143781	Hs.126280	hypothetical protein FLJ23393	14.00		
	114736	AI610347	Hs.103812	ESTs, Moderately similar to ALU1_HUMAN A		4.20	
	114596	AA310162	Hs.169248	cytochrome c	10.71		
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	20.40		
35	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.40		
	114452	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE		17.20	
	114359	NM_016929	Hs.283021	chloride intracellular channel 5		2.09	
	114357	R41677	Hs.6107	Homo sapiens cDNA FLJ14839 fis, clone OV	12.40		
	114251	H15261	Hs.21948	ESTs		2.00	
40	114138	AW384793	Hs.15740	Homo sapiens mRNA; cDNA DKFZp434E033 (fr		11.40	
	114124	W57554	Hs.125019	ESTs		6.04	
	113946	AW083883	Hs.37896	Homo sapiens cDNA FLJ13510 fis, clone PL		1.82	
	113695	T96965	Hs.17948	ESTs, Weakly similar to ALUB_HUMAN !!!!			
	113606	NM_013343	Hs.278951	NAG-7 protein		2.15	
45	113590	R49642	Hs.142447	ESTs, Weakly similar to ALU1_HUMAN ALU S		3.60	
	113560	T91015	Hs.268626	ESTs	32.00		
	113552	AI654223	Hs.16026	hypothetical protein FLJ23191			
	113540	AW152618	Hs.16757	ESTs			
	113502	T89130		gb:ye12d01.s1 Stratagene lung (937210) H		8.35	
50	113288	AI076838	Hs.12967	ESTs	12.40		
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en		4.27	
	113238	R45467	Hs.189813	ESTs			
	113203	AA743563	Hs.10305	ESTs	21.20		
	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom		1.92	
55	113089	T40707	Hs.270862	ESTs	14.33		
	113076	AF033199	Hs.8198	zinc finger protein 204		6.00	
	113009	T23699	Hs.7246	ESTs		9.40	
	112937	AI694320	Hs.6295	ESTs, Weakly similar to T17248 hypotheti		12.20	
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-li	10.57		
60	112794	R97018		gb:yq74b08.s1 Soares fetal liver spleen	26.60		
	112691	R88708	Hs.220647	ESTs	15.33		
	112602	AW004045	Hs.203365	ESTs	15.60		
	112366	AF035318	Hs.12533	Homo sapiens clone 23705 mRNA sequence	15.40		
	112210	R49645	Hs.7004	ESTs	14.00		
65	112064	AL049390	Hs.22689	Homo sapiens mRNA; cDNA DKFZp586O1318 (f	13.00		
	111998	R42379	Hs.138283	ESTs	11.00		
	111987	NM_015310	Hs.6763	KIAA0942 protein	22.40		
	111803	AA593731	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A		1.77	
	111737	H04607	Hs.9218	ESTs		1.86	
70	111605	T91061	Hs.194178	ESTs, Moderately similar to PC4259 ferri	23.00		
	111510	R07856	Hs.16355	ESTs	11.02		
	111341	AL157484	Hs.22483	Homo sapiens mRNA; cDNA DKFZp762M127 (fr		1.88	
	111280	AA373527	Hs.19385	CGI-58 protein	18.40		
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f			
75	111232	AI247763	Hs.16928	ESTs	27.60		
	110942	R63503	Hs.28419	ESTs	14.80		
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	24.71		
	110837	H03109	Hs.108920	HT018 protein		2.18	
	110824	AI767183	Hs.26942	ESTs	12.20		
80	110776	AB032417	Hs.19545	frizzled (Drosophila) homolog 4		1.75	
	110576	H60869	Hs.37889	ESTs	13.00		
	110369	AK000768	Hs.107872	hypothetical protein FLJ20761		5.60	
	110099	R44557	Hs.23748	ESTs		2.31	
	109984	AI796320	Hs.10299	Homo sapiens cDNA FLJ13545 fis, clone PL			
85	109958	AA001266	Hs.133521	ESTs	11.25		
	109893	AA884208	Hs.30484	ESTs		2.68	



5	109842	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	23.83		
	109837	H00656	Hs.29792	ESTs, Weakly similar to I38022 hypotheti		17.20	3.91
	109796	AI800515	Hs.12024	ESTs		9.60	
	109688	R41900	Hs.22245	ESTs			
	109648	H17800	Hs.7154	ESTs	22.80		
10	109613	H47315	Hs.27519	ESTs			
	109550	AW021488	Hs.26981	ESTs			
	109523	AW193342	Hs.24144	ESTs			1.89
	109472	AK001989	Hs.91165	hypothetical protein		6.00	
	109355	AA524525	Hs.48297	DKFZP586C1620 protein	15.00		
15	109260	AW978515	Hs.131915	KIAA0863 protein	25.60		
	108781	AA128654		gb:zn98g07.s1 Stratagene fetal retina 93	14.20		
	108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypotheti	11.00		
	108573	AA086005		gb:zl84c04.s1 Stratagene colon (937204)	26.00		
	108480	AL133092	Hs.68055	hypothetical protein DKFZp434I0428			
20	108382	NM_006770	Hs.67726	macrophage receptor with collagenous str			1.83
	108174	AA055632	Hs.303070	ESTs	15.20		
	108138	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp564G112 (fr		3.60	
	108087	AA045708	Hs.40545	ESTs	15.44		
	108048	AI797341	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT		11.40	
25	108041	AW204712	Hs.61957	ESTs			
	107997	AL049176	Hs.82223	chordin-like		4.76	
	107994	AA036811	Hs.48469	LIM domains containing 1			
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	14.20		
	107681	BE379594	Hs.49136	ESTs, Moderately similar to ALU7_HUMAN A	51.80		
30	107666	AA010611	Hs.60418	EST	29.20		
	107332	T87750	Hs.183297	DKFZP566F2124 protein	10.73		
	107292	BE166479	Hs.4789	Homo sapiens serologically defined breas	32.00		
	107230	AI034467	Hs.34650	ESTs	17.40		
	107168	W57578	Hs.237955	RAB7, member RAS oncogene family	10.43		
35	107160	AA314490	Hs.27669	KIAA1563 protein	11.40		
	107054	AI076459	Hs.15978	KIAA1272 protein			
	107029	AF264750	Hs.288971	myeloid/lymphoid or mixed-lineage leukem	21.40		
	106999	H93281	Hs.10710	hypothetical protein FLJ20417	35.80		
	106954	AF128847	Hs.204038	indolethylamine N-methyltransferase			1.76
40	106870	AI983730	Hs.26530	serum deprivation response (phosphatidyl			
	106865	AW192535	Hs.19479	ESTs	13.40		
	106844	AA485055	Hs.158213	sperm associated antigen 6		7.13	
	106820	NM_016831	Hs.12592	period (Drosophila) homolog 3		7.00	
	106818	AK002135	Hs.3542	hypothetical protein FLJ11273	13.00		
45	106797	AI768801	Hs.169943	Homo sapiens cDNA FLJ13569 fis, clone PL			2.05
	106773	AA478109	Hs.188833	ESTs			
	106747	NM_007118	Hs.171957	triple functional domain (PTPRF interact	12.60		
	106743	BE613328	Hs.21938	hypothetical protein FLJ12492	10.60		
	106667	AW360847	Hs.16578	ESTs			
50	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr			2.40
	106567	AW450408	Hs.86412	chromosome 9 open reading frame 5			1.78
	106562	AL031846	Hs.152151	plakophilin 4			1.76
	106536	AA329648	Hs.23804	ESTs, Weakly similar to PN0099 son3 prot			2.19
	106533	AL134708	Hs.145998	ESTs	23.20		
55	106507	AA259068	Hs.267819	protein phosphatase 1, regulatory (inhib	15.20		
	106490	AA404265	Hs.115537	putative dipeptidase			
	106474	BE383668	Hs.42484	hypothetical protein FLJ10618	10.44		
	106211	AA428240	Hs.126083	ESTs		29.80	
	105986	AB037722	Hs.8707	KIAA1301 protein		3.70	
60	105894	AI904740	Hs.25691	receptor (calcitonin) activity modifying			1.94
	105847	AW964490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-			1.75
	105803	AW747996	Hs.160999	ESTs, Moderately similar to A56194 throm			2.47
	105731	AA834664	Hs.29131	nuclear receptor coactivator 2	10.71		
	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds			
65	105688	AI299139	Hs.17517	ESTs	23.40		
	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	37.20		
	105101	H63202	Hs.38163	ESTs		8.30	
	104989	R65998	Hs.285243	hypothetical protein FLJ22029		8.09	
	104986	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1			1.92
70	104969	AI670947	Hs.78406	phosphatidylinositol-4-phosphate 5-kinas		5.40	
	104903	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,		7.60	
	104896	AW015318	Hs.23165	ESTs	13.80		
	104865	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C			
	104825	AA035613	Hs.141883	ESTs			1.87
75	104781	AA099904	Hs.21610	DKFZP434B203 protein			1.93
	104776	AA026349		gb:zj99f01.s1 Soares_pregnant_uterus_NbH		10.20	
	104691	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor		5.69	
	104667	AI239923	Hs.30098	ESTs		3.82	
	104404	H58762		gb:EST00057 HE6W Homo sapiens cDNA clone		4.20	
80	104392	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	27.20		
	104212	AB002298	Hs.173035	KIAA0300 protein			1.91
	104074	AL162039	Hs.31422	Homo sapiens mRNA; cDNA DKFZp434M229 (fr	11.20		
	103749	AL135301	Hs.8768	hypothetical protein FLJ10849	10.86		
	103645	AW246253	Hs.7043	succinate-CoA ligase, GDP-forming, alpha	12.00		
85	103554	AI878826	Hs.323469	caveolin 1, caveolae protein, 22kD			1.80
	103541	AI815601	Hs.79197	CD83 antigen (activated B lymphocytes, i			
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2			
	103428	BE383507	Hs.78921	A kinase (PRKA) anchor protein 1	11.20		
	103353	X89399	Hs.119274	RAS p21 protein activator (GTPase activa	19.80		



	103295	X81479	Hs.2375	egf-like module containing, mucin-like,	3.60	
	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula		
	103100	NM_005574	Hs.184585	LIM domain only 2 (rhombotin-like 1)		1.76
5	103025	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t		2.15
	102698	M18667	Hs.1867	progastricsin (pepsinogen C)		
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein	11.00	
	102580	U60808	Hs.152981	CDP-diacylglycerol synthase (phosphatida	25.40	
	102417	AA034127	Hs.153487	signal transducing adaptor molecule (SH3	14.00	
10	102363	NM_003734	Hs.198241	amine oxidase, copper containing 3 (vasc		
	102302	AA306342	Hs.69171	protein kinase C-like 2	10.86	
	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11		
	102188	U20350	Hs.78913	chemokine (C-X3-C) receptor 1		7.40
	102151	T27013	Hs.3132	steroidogenic acute regulatory protein	16.40	
15	101957	L28824	Hs.74101	spleen tyrosine kinase	15.40	
	101842	M93221	Hs.75182	mannose receptor, C type 1		
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant		
	101764	AI198550	Hs.81256	S100 calcium-binding protein A4 (calcium		1.78
	101716	AF050658	Hs.2563	tachykinin, precursor 1 (substance K, su	18.80	
20	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a I		2.22
	101447	M21305		gb:Human alpha satellite and satellite 3	504.80	
	101383	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co		31.00
	101346	AI738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N		1.75
	101345	NM_005795	Hs.152175	calcitonin receptor-like		
25	101336	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h		2.24
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do		
	101277	BE297626	Hs.296049	microfibrillar-associated protein 4		
	101262	L35854		gb:Human dystrophin (dp140) mRNA, 5' end	19.00	
	101168	NM_005308	Hs.211569	G protein-coupled receptor kinase 5		2.01
30	101102	NM_003243	Hs.79059	transforming growth factor, beta recepto		
	101088	X70697	Hs.553	solute carrier family 6 (neurotransmitte		7.52
	101066	AW970254	Hs.889	Charot-Leyden crystal protein	19.38	
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		1.91
	100893	BE245294	Hs.180789	S164 protein	15.40	
35	100770	W25797.comp	Hs.177486	amyloid beta (A4) precursor protein (pro	11.20	
	100716	X89887	Hs.172350	HIR (histone cell cycle regulation defec	14.80	
	100555	M69181		gb:Human nonmuscle myosin heavy chain-B	33.00	
	100425	NM_014747	Hs.78748	KIAA0237 gene product	16.20	
	100408	D86640	Hs.56045	src homology three (SH3) and cysteine ri		4.00
40	100382	D83407	Hs.156007	Down syndrome critical region gene 1-lik		4.24
	100351	D64158				6.20
	100299	D49493	Hs.2171	growth differentiation factor 10		21.20
	100134	AA305746	Hs.49	macrophage scavenger receptor 1		
	100108	U09577	Hs.76873	hyaluronoglucosaminidase 2		1.79
45	100095	Z97171	Hs.78454	myocilin, trabecular meshwork inducible		5.40
	100066				11.29	

TABLE 3B shows the accession numbers for those primekeys lacking unigenelD's for Table 3A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using *Clustering and Alignment Tools (DoubleTwist, Oakland California)*. The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

60	Pkey	CAT number	Accessions
	123619	371681_1	AA602964 AA609200
	126433	127143_1	AA325606 AA099517 N89423
	125831	1522905_1	H04043 D60988 D60337
65	126816	122973_1	AA248234 AA090985
	126852	136135_1	AA399961 AA128347
	121059	273450_1	AA393283 AA398628
	120637	200885_1	AA811804 AA809404 AA286907 AW977624
	122011	7617_-2	AA431082
70	120934	177521_1	AA226198 AA226513 AA383773
	123802	genbank_AA620448	AA620448
	116814	genbank_H50834	H50834
	118329	genbank_N63520	N63520
	104404	H58762_at	H58762
75	104776	genbank_AA026349	AA026349
	113502	genbank_T89130T89130	
	101262	entrez_L35854	L35854
	108573	genbank_AA086005	AA086005
	101447	entrez_M21305	M21305
80	124357	genbank_N22401	N22401
	108781	genbank_AA128654	AA128654
	112794	genbank_R97018	R97018
	100351	entrez_D64158	D64158
85	100555	tigr_HT2245	M69181 M81105 U51039



Table 4A shows 202 genes up-regulated in samples from patients treated with chemotherapy or radiotherapy. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number		
	ExAccn:	Exemplar Accession number, Genbank accession number		
	UnigenelD:	Unigene number		
	Unigene Title:	Unigene gene title		
10	R1:	average of AI for samples from patients treated with chemotherapy or radiotherapy divided by the average of AI for normal lung samples.		
	Pkey	ExAccn	UnigenelD	Unigene Title
				R1
	100113	NM_001269	Hs.84746	chromosome condensation 1
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3
15	100210	D26361	Hs.3104	KIAA0042 gene product
	100225	D28539	Hs.167185	glutamate receptor, metabotropic 5
	100269	NM_001949	Hs.1189	E2F transcription factor 3
	100438	AA013051	Hs.91417	topoisomerase (DNA) II binding protein
	100877	X80821	Hs.27973	KIAA0874 protein
20	100893	BE245294	Hs.180789	S164 protein
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto
	101447	M21305		gb:Human alpha satellite and satellite 3
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr
	101724	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)
25	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn
	101879	AA176374	Hs.243886	nuclear autoantigenic sperm protein (his
	101915	AF207881	Hs.155185	cytosolic ovarian carcinoma antigen 1
	101973	U41514	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp
30	102025	U04045	Hs.78934	mutS (E. coli) homolog 2 (colon cancer,
	102031	U04898	Hs.2156	RAR-related orphan receptor A
	102052	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma
	102391	AA296874	Hs.77494	deoxyguanosine kinase
	102420	U44060	Hs.14427	Homo sapiens cDNA: FLJ21800 fis, clone H
35	102610	U65011	Hs.30743	preferentially expressed antigen in mela
	102829	NM_006183	Hs.80962	neurotensin
	103000	NM_001975	Hs.146580	enolase 2, (gamma, neuronal)
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	103507	AJ000512	Hs.296323	serum/glucocorticoid regulated kinase
40	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
	104660	BE298665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr
	104896	AW015318	Hs.23165	ESTs
	105038	AW503733	Hs.9414	KIAA1488 protein
	105298	BE387790	Hs.26369	hypothetical protein FLJ20287
45	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specif
	106073	AL157441	Hs.17834	downstream neighbor of SON
	106205	AW965058	Hs.111583	ESTs, Weakly similar to I38022 hypotheti
	106516	AL137311	Hs.234074	Homo sapiens mRNA; cDNA DKFZp761G02121 (
50	106533	AL134708	Hs.145998	ESTs
	106575	AW970602	Hs.105421	ESTs
	106654	AW075485	Hs.286049	phosphoserine aminotransferase
	106851	AI458623		gb:tk04g09.x1 NCI_CGAP_Lu24 Homo sapiens
	106995	AB023139	Hs.37892	KIAA0922 protein
55	107332	T87750	Hs.183297	DKFZP566F2124 protein
	107532	AA443473	Hs.173684	Homo sapiens mRNA; cDNA DKFZp762G207 (fr
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR
	108609	BE409857	Hs.69499	hypothetical protein
	108780	AU076442	Hs.117938	collagen, type XVII, alpha 1
60	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines
	109260	AW978515	Hs.131915	KIAA0863 protein
	109280	AK001355	Hs.279610	hypothetical protein FLJ10493
	109292	AW975746	Hs.188662	KIAA1702 protein
	109384	AA219172	Hs.86849	ESTs
65	109415	U80736	Hs.110826	trinucleotide repeat containing 9
	109445	AA232103	Hs.189915	ESTs
	109502	AW967069	Hs.211556	hypothetical protein MGC5487
	109633	AW003785	Hs.170267	ESTs
	109786	AI989482	Hs.146286	kinesin family member 13A
70	109958	AA001266	Hs.133521	ESTs
	110920	N47224	Hs.20521	HMT1 (hnRNP methyltransferase, S. cerevi
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1
	111084	H44186	Hs.15456	PDZ domain containing 1
	111132	AB037807	Hs.83293	hypothetical protein
75	111229	AW389845	Hs.110855	ESTs
	111337	AA837396	Hs.263925	LIS1-interacting protein NUDE1, rat homo
	111987	NM_015310	Hs.6763	KIAA0942 protein
	112046	AA383343	Hs.22116	CDC14 (cell division cycle 14, S. cerevi
	112268	W39609	Hs.22003	solute carrier family 6 (neurotransmitte
80	112685	R87650	Hs.33439	ESTs, Weakly similar to ALU1_HUMAN ALU
	112871	AL110216	Hs.12285	ESTs, Weakly similar to I55214 salivary
	112897	AW206453	Hs.3782	ESTs
	112973	AB033023	Hs.318127	hypothetical protein FLJ10201
	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f
85	113073	N39342	Hs.103042	microtubule-associated protein 1B



	113494	T91451	Hs.86538	ESTs	22.80
	113560	T91015	Hs.268626	ESTs	22.80
	113849	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	51.80
5	113950	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	28.20
	114339	AA782845	Hs.22790	ESTs	20.20
	114365	H42169	Hs.18653	hypothetical protein FLJ14627	21.00
	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	25.80
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	23.60
10	114824	AA960961	Hs.305953	zinc finger protein 83 (HPF1)	27.20
	114837	BE244930	Hs.166895	ESTs	30.20
	114974	AW966931	Hs.179662	nucleosome assembly protein 1-like 1	20.80
	115075	AA814043	Hs.88045	ESTs	30.60
	115084	BE383668	Hs.42484	hypothetical protein FLJ10618	28.86
15	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	38.00
	115313	AA808001	Hs.184411	albumin	22.60
	115697	D31382	Hs.63325	transmembrane protease, serine 4	173.60
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	27.77
	116090	AI591147	Hs.61232	ESTs	20.80
20	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	164.20
	116399	AA889120	Hs.110637	homeo box A10	38.00
	117099	H93699		gb:yv16a11.s1 Soares fetal liver spleen	21.60
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	49.40
	118091	AW005054	Hs.47883	ESTs, Weakly similar to KCC1_HUMAN CALCI	22.40
25	118138	AA374756	Hs.93560	Homo sapiens mRNA for KIAA1771 protein,	22.00
	118720	N73515		gb:za49d07.s1 Soares fetal liver spleen	20.00
	118873	AI824009	Hs.44577	ESTs	19.40
	119126	R45175	Hs.117183	ESTs	111.20
	119717	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	33.00
30	119940	AL050097	Hs.272531	DKFZP586B0319 protein	31.00
	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	20.20
	120515	AA258356		gb:zr59c10.s1 Soares_NhHMPu_S1 Homo sapi	25.00
	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	95.40
	120983	AA398209	Hs.97587	EST	105.20
35	121054	AW976570	Hs.97387	ESTs	38.80
	121369	AW450737	Hs.128791	CGI-09 protein	41.60
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	30.80
	122612	AA974832	Hs.128708	ESTs	19.60
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H	33.20
40	123440	AI733692	Hs.112488	ESTs	23.17
	123596	AA421130	Hs.112640	EST	23.00
	123619	AA602964		gb:no97c02.s1 NCI_CGAP_Pr2 Homo sapiens	28.80
	124006	AI147155	Hs.270016	ESTs	77.60
	124169	BE079334	Hs.271630	ESTs	22.20
45	124281	AI333756	Hs.111801	arsenate resistance protein ARS2	42.20
	124472	N52517	Hs.102670	EST	32.60
	124617	AW628168	Hs.152684	ESTs	21.80
	124631	NM_014053	Hs.270594	FLVCR protein	30.40
	124839	R55784	Hs.140942	ESTs	21.20
50	125186	AA610620	Hs.181244	major histocompatibility complex, class	42.80
	125321	T86652	Hs.178294	ESTs	27.00
	125535	NM_013243	Hs.22215	secretogranin III	23.80
	125646	AA628962	Hs.75209	protein kinase (cAMP-dependent, catalyti	23.20
	125684	AW589427	Hs.158849	Homo sapiens cDNA: FLJ21663 fis, clone C	21.20
55	125724	AL360190	Hs.295978	Homo sapiens mRNA full length insert cDN	48.80
	125847	AW161885	Hs.249034	ESTs	31.00
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	21.20
	126077	M78772	Hs.210836	ESTs	49.80
	126299	AW979155	Hs.298275	amino acid transporter 2	21.80
60	126395	AI468004	Hs.278956	hypothetical protein FLJ12929	71.00
	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	23.20
	126509	R47400	Hs.23850	ESTs	23.80
	126538	AB030656	Hs.17377	coronin, actin-binding protein, 1C	23.10
	126666	AA648886	Hs.151999	ESTs	36.00
65	126812	AB037860	Hs.173933	nuclear factor I/A	20.80
	126872	AW450979		gb:UI-H-BI3-ala-a-12-0-UI.s1 NCI_CGAP_Su	46.29
	127046	AA321948	Hs.293968	ESTs	22.80
	127431	AW771958	Hs.175437	ESTs, Moderately similar to PC4259 ferri	30.00
	127489	AA650250	Hs.272076	ESTs	20.80
70	127521	AW297206	Hs.164018	ESTs	25.20
	127742	AW293496	Hs.180138	ESTs	28.00
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	21.20
	127930	AA809672	Hs.123304	ESTs	20.54
	127968	AA830201	Hs.124347	ESTs	28.20
75	127987	AI022103	Hs.124511	ESTs	19.60
	128116	H07103	Hs.286014	Homo sapiens, clone IMAGE:3867243, mRNA	20.40
	128609	NM_003616	Hs.102456	survival of motor neuron protein interac	34.40
	128777	AI878918	Hs.10526	cysteine and glycine-rich protein 2	53.80
	128949	AA009647	Hs.8850	a disintegrin and metalloproteinase doma	23.00
80	129168	AI132988	Hs.109052	chromosome 14 open reading frame 2	37.60
	129404	AI267700	Hs.317584	ESTs	28.60
	129527	AA769221	Hs.270847	delta-tubulin	40.80
	129574	AA026815	Hs.11463	UMP-CMP kinase	31.20
	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	29.60
	129785	H19006	Hs.184780	ESTs	72.20
85	129970	AV655806	Hs.296198	chromosome 12 open reading frame 4	22.20



	130149	AW067805	Hs.172665	methylenetetrahydrofolate dehydrogenase	29.60
	130199	Z48579	Hs.172028	a disintegrin and metalloproteinase doma	27.60
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	28.36
5	130466	W19744	Hs.180059	Homo sapiens cDNA FLJ20653 fis, clone KA	20.20
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	22.40
	130617	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	19.60
	130703	R77776	Hs.18103	ESTs	19.40
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	21.40
10	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	110.00
	131028	AI879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	25.20
	131086	AL035461	Hs.2281	chromogranin B (secretogranin 1)	40.60
	131284	NM_001429	Hs.25272	E1A binding protein p300	24.60
	131775	AB014548	Hs.31921	KIAA0648 protein	21.00
15	131860	BE383676	Hs.334	Rho guanine nucleotide exchange factor (	33.40
	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	60.80
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	20.40
	132084	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)	29.40
	132389	AA310393	Hs.190044	ESTs	32.40
20	132437	AA152106	Hs.4859	cyclin L ania-6a	27.40
	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	75.60
	132617	AF037335	Hs.5338	carbonic anhydrase XII	31.36
	132632	AU076916	Hs.5398	guanine monophosphate synthetase	32.40
	132672	W27721	Hs.54697	Cdc42 guanine exchange factor (GEF) 9	23.40
25	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypothei	61.20
	132771	Y10275	Hs.56407	phosphoserine phosphatase	22.33
	133070	U92649	Hs.64311	a disintegrin and metalloproteinase doma	23.50
	133153	AF070592	Hs.66170	HSKM-B protein	30.00
	133181	X91662	Hs.66744	twist ( <i>Drosophila</i> ) homolog ( <i>acrocephalos</i>	23.80
30	133282	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	51.60
	133350	AI499220	Hs.71573	hypothetical protein FLJ10074	33.00
	133592	AV652066	Hs.75113	general transcription factor IIIA	82.00
	133658	AA319146	Hs.75426	secretogranin II (chromogranin C)	
	133865	AB011155	Hs.170290	discs, large ( <i>Drosophila</i> ) homolog 5	69.33
35	134032	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibito	33.20
	134125	NM_014781	Hs.50421	KIAA0203 gene product	31.60
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	30.60
	134321	BE538082	Hs.8172	ESTs, Moderately similar to A46010 X-lin	23.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	49.20
40	134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	20.20
	134753	NM_006482	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl	20.80
	135002	AA448542	Hs.251677	G antigen 7B	37.60
	135029	H58818	Hs.187579	hydroxysteroid (17-beta) dehydrogenase	53.40
	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	31.60
45	135345	X53655	Hs.99171	neurotrophin 3	28.80

TABLE 4B shows the accession numbers for those primekeys lacking unigenelD's for Table 4A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
123619	371681_1	AA602964 AA609200
126433	127143_1	AA325606 AA099517 N89423
126872	142696_1	AW450979 AA136653 AA136656 AW419381 AA984358 AA492073 BE168945 AA809054 AW238038 BE011212 BE011359 BE011367 BE011368 BE011362 BE011215 BE011365 BE011363
106851	322947_1	AI458623 AA639708 AA485409 R22065 AA485570
118720	genbank_N73515	N73515
120515	genbank_AA258356	AA258356
117099	321871_1	H93699 H97976 H80036
101447	entrez_M21305	M21305
123130	genbank_AA487200	AA487200



Table 5A shows 680 genes up-regulated in squamous cell carcinoma or adenocarcinoma lung tumors relative to normal lung and chronically diseased lung. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number							
	ExAccn:	Exemplar Accession number, Genbank accession number							
	UnigeneID:	Unigene number							
	Unigene Title:	Unigene gene title							
10	R1:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R2:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R3:	80th percentile of AI squamous cell carcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R4:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 80th percentile of AI for squamous cell carcinoma lung tumor samples.							
15	R5:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by 90th percentile of AI for normal and chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples							
20	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5
	100035			AFFX control: GAPDH					6.76
	100036			AFFX control: GAPDH					5.77
	100037			AFFX control: GAPDH					5.75
25	100071	A28102		Human GABA <sub>A</sub> receptor alpha-3 subunit		8.00			
	100114	X02308	Hs.82962	thymidylate synthetase					5.71
	100154	H60720	Hs.81892	KIAA0101 gene product	3.84				
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	3.33				
	100188	AW247090	Hs.57101	minichromosome maintenance deficient (S.					4.52
	100202	BE294407	Hs.99910	phosphofructokinase, platelet					5.49
30	100216	AA489908	Hs.1390	proteasome (prosome, macropain) subunit,					5.67
	100269	NM_001949	Hs.1189	E2F transcription factor 3	2.55				
	100287	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e					5.66
	100297	AU077258	Hs.182429	protein disulfide isomerase-related prot					3.81
	100330	AW410976	Hs.77152	minichromosome maintenance deficient (S.					4.50
35	100335	AW247529	Hs.6793	platelet-activating factor acetylhydrola	5.07				
	100360	W70171	Hs.75939	uridine monophosphate kinase					4.82
	100372	NM_014791	Hs.184339	KIAA0175 gene product					3.79
	100474	NM_000699	Hs.300280	amylase, alpha 2A; pancreatic				15.65	
	100486	T19006	Hs.10842	RAN, member RAS oncogene family					5.49
40	100491	D56165	Hs.275163	non-metastatic cells 2, protein (NM23B)					4.17
	100516	D90278	Hs.11	carcinoembryonic antigen-related cell ad		7.20			
	100522	X51501	Hs.99949	prolactin-induced protein				14.20	
	100559	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys	3.10				
	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid				9.30	
45	100629	AA015693	Hs.21291	mitogen-activated protein kinase kinase				20.60	
	100661	BE623001	Hs.132748	Homo sapiens ribosomal protein L39 mRNA,	3.85				
	100677	AA353686	Hs.57813	zinc ribbon domain containing, 1		8.60			
	100696	D14887	Hs.121686	general transcription factor IIA, 1 (37k				10.00	
	100709	N26539	Hs.100469	myeloid/lymphoid or mixed-lineage leukem			24.80		
50	100761	BE208491	Hs.295112	KIAA0618 gene product		7.60			
	100830	AC004770	Hs.4756	flap structure-specific endonuclease 1					7.99
	100867	U14622		gb:Human transketolase-like protein gene		10.20			
	100902	M16029	Hs.287270	ret proto-oncogene (multiple endocrine n		8.00			
	100906	AU076916	Hs.5398	guanine monphosphate synthetase					5.16
55	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	2.57				
	101045	J05614		gb:Human proliferating cell nuclear anti					4.69
	101061	NM_000175	Hs.180532	glucose phosphate isomerase					4.19
	101071	L02840	Hs.84244	potassium voltage-gated channel, Shab-re		12.91			
	101124	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	3.12				
60	101175	U82671	Hs.36980	melanoma antigen, family A, 2	3.50				
	101181	BE262621	Hs.73798	macrophage migration inhibitory factor (					5.69
	101204	L24203	Hs.82237	ataxia-telangiectasia group D-associated	4.08				
	101210	L29301	Hs.2353	opioid receptor, mu 1			6.40		
	101216	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	2.53				
65	101228	AA333387	Hs.82916	chaperonin containing TCP1, subunit 6A (					7.90
	101233	AL135173	Hs.878	sorbitol dehydrogenase					4.45
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	8.50				
	101342	U52112	Hs.182018	interleukin-1 receptor-associated kinase					4.17
	101346	A1738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				21.89	
70	101369	NM_000892	Hs.1901	kallikrein B, plasma (Fletcher factor) 1				12.80	
	101396	BE267931	Hs.78996	proliferating cell nuclear antigen	3.24				
	101431	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)					7.90
	101448	NM_000424	Hs.195850	keratin 5 (epidermolysis bullosa simplex	8.31				
	101462	AL035668	Hs.73853	bone morphogenetic protein 2				38.80	
75	101466	BE262660	Hs.170197	glutamic-oxaloacetic transaminase 2, mit					4.01
	101484	AA053486	Hs.20315	interferon-induced protein with tetratri				12.00	
	101502	M26958		gb:Human parathyroid hormone-related pro	10.50				
	101505	AA307680	Hs.75692	asparagine synthetase					4.46
	101526	NM_002197	Hs.154721	aconitase 1, soluble	4.02				
80	101535	X57152	Hs.99853	fibrillarin					4.65
	101577	M34353	Hs.1041	v-ros avian UR2 sarcoma virus oncogene h				9.09	
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	54.00				
	101663	NM_003528	Hs.2178	H2B histone family, member Q	5.59				
	101664	AA436989	Hs.121017	H2A histone family, member A	7.00				
85	101669	L24498	Hs.80409	growth arrest and DNA-damage-inducible,		7.60			



	101695	M69136	Hs.135626	chymase 1, mast cell	4.79			
	101724	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	15.21			
	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	55.50			
5	101759	M80244	Hs.184601	solute carrier family 7 (cationic amino			4.10	
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant		18.57		
	101804	M86699	Hs.169840	TTK protein kinase	4.50			
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	140.00			
	101833	AU076442	Hs.117938	collagen, type XVII, alpha 1	2.56			
10	101842	M93221	Hs.75182	mannose receptor, C type 1		12.80		
	101851	BE260964	Hs.82045	midkine (neurite growth-promoting factor			5.88	
	102002	NM_002484	Hs.81469	nucleotide binding protein 1 (E.coli Min	7.80			
	102039	AL134223	Hs.306098	aldo-keto reductase family 1, member C1			4.35	
	102072	U09410	Hs.78743	zinc finger protein 131 (clone pHZ-10)		7.40		
15	102083	T35901	Hs.75117	interleukin enhancer binding factor 2, 4			5.12	
	102111	L36196	Hs.81884	sulfotransferase family, cytosolic, 2A,		12.00		
	102123	NM_001809	Hs.1594	centromere protein A (17kD)	6.20			
	102154	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalinin	2.62			
	102193	AL036335	Hs.313	secreted phosphoprotein 1 (osteopontin,	5.85			
20	102217	AA829978	Hs.301613	JTV1 gene			6.18	
	102224	NM_002810	Hs.148495	proteasome (prosome, macropain) 26S subu			4.49	
	102234	AW163390	Hs.278554	heterochromatin-like protein 1			5.80	
	102251	NM_004398	Hs.41706	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	4.50			
	102305	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)			5.15	
25	102330	BE298063	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta			4.17	
	102340	U37055	Hs.278657	macrophage stimulating 1 (hepatocyte gro		9.33		
	102348	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	8.87			
	102368	U39817	Hs.36820	Bloom syndrome	15.91			
	102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma		19.20		
30	102404	NM_005429	Hs.79141	vascular endothelial growth factor C			14.00	
	102537	U57094	Hs.50477	RAB27A, member RAS oncogene family			12.00	
	102581	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2			4.57	
	102605	AI435128	Hs.181369	ubiquitin fusion degradation 1-like			3.98	
	102610	U65011	Hs.30743	preferentially expressed antigen in mela	77.50			
35	102623	AW249285	Hs.37110	melanoma antigen, family A, 9	12.50			
	102642	AA205847	Hs.23016	G protein-coupled receptor		22.00		
	102654	AV649989	Hs.24385	Human hbc647 mRNA sequence		12.00		
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein			12.80	
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	6.50			
40	102672	U72066	Hs.29287	retinoblastoma-binding protein 8	8.50			
	102687	NM_007019	Hs.93002	ubiquitin carrier protein E2-C			9.24	
	102696	BE540274	Hs.239	forkhead box M1			5.54	
	102768	U82321		gb:Homo sapiens clone 14.9B mRNA sequenc		6.60		
	102781	BE258778	Hs.108809	chaperonin containing TCP1, subunit 7 (e			3.78	
45	102784	U85658	Hs.61796	transcription factor AP-2 gamma (activat			4.26	
	102824	U90916	Hs.82845	Homo sapiens cDNA: FLJ21930 fis, clone H		14.40		
	102829	NM_006183	Hs.80962	neurotensin	8.00			
	102888	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1			5.50	
	102892	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin		6.70		
50	102913	NM_002275	Hs.80342	keratin 15	4.64			
	102935	BE561850	Hs.80506	small nuclear ribonucleoprotein polypept	2.93			
	102951	X15218	Hs.2969	v-ski avian sarcoma viral oncogene homol			11.40	
	102983	BE387202	Hs.118638	non-metastatic cells 1, protein (NM23A)			7.26	
	103023	AW500470	Hs.117950	multifunctional polypeptide similar to S	3.01			
55	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	27.90			
	103038	AA926960	Hs.334883	CDC28 protein kinase 1			8.79	
	103060	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin			4.27	
	103099	AI693251	Hs.8248	NADH dehydrogenase (ubiquinone) Fe-S pro		9.80		
	103119	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	4.05			
60	103168	X53463	Hs.2704	glutathione peroxidase 2 (gastrointestin	3.07			
	103185	NM_006825	Hs.74368	transmembrane protein (63kD), endoplasmic			5.62	
	103192	M22440	Hs.170009	transforming growth factor, alpha		7.40		
	103223	BE275607	Hs.1708	chaperonin containing TCP1, subunit 3 (g			4.70	
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o		100.00		
65	103316	X83301	Hs.324728	SMA5			9.80	
	103375	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	9.71			
	103376	AL036166	Hs.323378	coated vesicle membrane protein	14.00			
	103385	NM_007069	Hs.37189	similar to rat HREV107			11.00	
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	2.93			
70	103404	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,			5.15	
	103430	BE564090	Hs.20716	translocase of inner mitochondrial membr			3.98	
	103446	X98834	Hs.79971	sal (Drosophila)-like 2			21.40	
	103476	Y07701	Hs.293007	aminopeptidase puromycin sensitive		13.00		
	103477	AJ011812	Hs.119018	transcription factor NRF		6.40		
75	103478	BE514982	Hs.38991	S100 calcium-binding protein A2	5.02			
	103515	Y10275	Hs.56407	phosphoserine phosphatase	10.50			
	103558	BE616547	Hs.2785	keratin 17	6.41			
	103580	AA328046	Hs.46405	polymerase (RNA) II (DNA directed) polyp			3.84	
	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	78.50			
80	103594	AI368680	Hs.816	SRY (sex determining region Y)-box 2	6.51			
	103636	NM_006235	Hs.2407	POU domain, class 2, associating factor	3.50			
	103768	AF086009		gb:Homo sapiens full length insert cDNA			4.48	
	103841	AA314821	Hs.38178	hypothetical protein FLJ23468		8.00		
	103847	AF219946	Hs.102237	tubby super-family protein		10.40		
85	103913	AW967500	Hs.133543	ESTs			15.60	
	104094	AA418187	Hs.330515	ESTs		6.60		



	104150	AL122044	Hs.331633	hypothetical protein DKFZp566N034		26.00	
	104257	BE560621	Hs.9222	estrogen receptor binding site associate	6.80		
	104261	AW248364	Hs.5409	RNA polymerase I subunit			3.98
5	104331	AB040450	Hs.279862	cdk inhibitor p21 binding protein	6.80		
	104415	BE410992	Hs.258730	heme-regulated initiation factor 2-alpha	10.29		
	104558	R56678	Hs.88959	hypothetical protein MGC4816	4.21		
	104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m		15.79	
	104658	AA360954	Hs.27268	Homo sapiens cDNA: FLJ21933 fis, clone H		17.40	
10	104660	BE298665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	6.40		
	104689	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr			6.55
	104754	AI206234	Hs.155924	cAMP responsive element modulator		10.00	
	104758	BE560269	Hs.7010	NPD002 protein			4.47
	104971	BE311926	Hs.15830	hypothetical protein FLJ12691	2.87		
15	105011	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	3.83		
	105012	AF098158	Hs.9329	chromosome 20 open reading frame 1	2.86		
	105026	AA809485	Hs.124219	hypothetical protein FLJ12934		11.00	
	105076	AI598252	Hs.37810	hypothetical protein MGC14833			5.01
	105132	AA148164	Hs.247280	HBV associated factor			3.99
20	105143	AI368836	Hs.24808	ESTs, Weakly similar to I38022 hypotheti		11.00	
	105158	AW976357	Hs.234545	hypothetical protein NUF2R		16.00	
	105175	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	4.32		
	105200	AA328102	Hs.24641	cytoskeleton associated protein 2	3.00		
	105264	AA227934		gb:zr57e08.s1 Soares_NhHMPu_S1 Homo sapi		10.00	
25	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	3.69		
	105409	AW505076	Hs.301855	DiGeorge syndrome critical region gene 8		9.20	
	105460	AW296078	Hs.271721	Homo sapiens, clone IMAGE:4179986, mRNA,		7.80	
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specif	4.12		
	105743	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	3.82		
30	105782	H09748	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro		27.00	
	105848	AW954064	Hs.24951	ESTs		7.60	
	105891	U55984	Hs.289088	heat shock 90kD protein 1, alpha			4.14
	106019	AF221993	Hs.46743	McKusick-Kaufman syndrome		16.80	
	106069	BE566623	Hs.29899	ESTs, Weakly similar to G02075 transcrip		23.40	
35	106073	AL157441	Hs.17834	downstream neighbor of SON	9.50		
	106126	AA576953	Hs.22972	hypothetical protein FLJ13352	6.00		
	106159	AK001301	Hs.3487	hypothetical protein FLJ10439			3.95
	106220	D61329	Hs.32196	mitochondrial ribosomal protein L36			6.04
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S		13.20	
40	106300	Y10043	Hs.19114	high-mobility group (nonhistone chromoso			5.02
	106307	AA436174	Hs.37751	ESTs, Weakly similar to putative p150 [	6.60		
	106318	AA025610	Hs.9605	cleavage and polyadenylation specific fa			5.04
	106341	AF191020	Hs.5243	hypothetical protein, estradiol-induced			7.25
	106440	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub		13.80	
45	106481	D61594	Hs.17279	tyrosylprotein sulfotransferase 1	4.75		
	106586	AA243837	Hs.57787	ESTs		10.84	
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr		45.60	
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	28.00		
	106785	Y15227	Hs.20149	deleted in lymphocytic leukemia, 1	3.00		
50	106813	C05766	Hs.181022	CGI-07 protein		11.40	
	106895	AK001826	Hs.25245	hypothetical protein FLJ11269		6.00	
	106913	AI219346	Hs.86178	M-phase phosphoprotein 9	6.56		
	106919	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S			4.27
	107054	AI076459	Hs.15978	KIAA1272 protein		34.80	
55	107059	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	4.71		
	107098	AI823593	Hs.27688	ESTs		24.80	
	107104	AU076640	Hs.15243	nucleolar protein 1 (120kD)			7.05
	107129	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.60		
	107198	AV657225	Hs.9846	KIAA1040 protein		19.20	
60	107203	D20426	Hs.41639	programmed cell death 2		7.60	
	107217	AL080235	Hs.35861	DKFZP586E1621 protein	9.50		
	107284	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	2.71		
	107318	T74445	Hs.5957	Homo sapiens clone 24416 mRNA sequence		8.71	
	107516	X57152	Hs.99853	fibrillarin			4.33
65	107529	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)			4.00
	107728	AA019551	Hs.294151	Homo sapiens, clone IMAGE:3603836, mRNA,	10.80		
	107851	AA022953	Hs.61172	EST		8.00	
	107901	L42612	Hs.335952	keratin 6B	3.40		
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	2.88		
70	107932	AW392555	Hs.18878	hypothetical protein FLJ21620	7.50		
	108015	AW298357	Hs.49927	protein kinase NYD-SP15		23.40	
	108056	AA043675	Hs.62633	ESTs		12.80	
	108075	AI867370	Hs.139709	hypothetical protein FLJ12572		12.80	
	108187	BE245374	Hs.27842	hypothetical protein FLJ11210	7.00		
75	108296	N31256	Hs.161623	ESTs	6.60		
	108305	AA071391		gb:zm61e06.r1 Stratagene fibroblast (937		11.80	
	108393	AA075211		gb:zm86a08.r1 Stratagene ovarian cancer		11.80	
	108480	AL133092	Hs.68055	hypothetical protein DKFZp434I0428		20.80	
	108554	AA084948		gb:zn13b09.s1 Stratagene hNT neuron (937	6.40		
80	108573	AA086005		gb:zl84c04.s1 Stratagene colon (937204)		25.40	
	108584	AA088326	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	9.60		
	108597	AK000292	Hs.278732	hypothetical protein FLJ20285		14.60	
	108695	AB029000	Hs.70823	KIAA1077 protein	3.00		
	108699	AA121514	Hs.70832	ESTs		10.00	
85	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c		11.00	
	108780	AU076442	Hs.117938	collagen, type XVII, alpha 1	11.21		



5	108810	AW295647	Hs.71331	hypothetical protein MGC5350	8.50				
	108816	AA130884	Hs.270501	ESTs, Moderately similar to ALU2_HUMAN		7.40			
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	4.00				
	108860	AA133334	Hs.129911	ESTs	6.09				
	108937	AL050107	Hs.24341	transcriptional co-activator with PDZ-bi	3.00				
10	109010	NM_007240	Hs.44229	dual specificity phosphatase 12	2.69				
	109121	BE389387	Hs.49767	NADH dehydrogenase (ubiquinone) Fe-S pro				4.53	
	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines	10.58				
	109227	AA766998	Hs.85874	Human DNA sequence from clone RP11-16L21		9.00			
	109415	U80736	Hs.110826	trinucleotide repeat containing 9		51.40			
15	109418	AI866946	Hs.161707	ESTs				11.00	
	109454	AA232255	Hs.295232	ESTs, Moderately similar to A46010 X-li			17.60		
	109502	AW967069	Hs.211556	hypothetical protein MGC5487			9.49		
	109543	AA564994	Hs.222851	ESTs		12.67			
	109648	H17800	Hs.7154	ESTs				10.40	
20	109680	AB037734	Hs.4993	KIAA1313 protein			33.20		
	109700	F09609		gb:HSC33H092 normalized infant brain cDN				16.00	
	109704	AI743880	Hs.12876	ESTs			11.00		
	109792	R49625		gb:yg61f03.s1 Soares infant brain 1NIB H				12.60	
	109981	BE546208	Hs.26090	hypothetical protein FLJ20272	4.00				
25	109998	AL042201	Hs.21273	transcription factor NYD-sp10		7.80			
	110039	H11938	Hs.21907	histone acetyltransferase		7.00			
	110156	AA581322	Hs.4213	hypothetical protein MGC16207				4.24	
	110500	AA907723	Hs.36962	ESTs	4.50				
	110551	AW450381	Hs.14529	ESTs		8.60			
30	110561	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.06				
	110854	BE612992	Hs.27931	hypothetical protein FLJ10607 similar to		6.80			
	110886	AW274992	Hs.72249	three-PDZ containing protein similar to			8.80		
	110916	BE178102	Hs.24349	ESTs		6.80			
	111003	N52980	Hs.83765	dihydrofolate reductase				16.80	
35	111337	AA837396	Hs.263925	LIS1-interacting protein NUDE1, rat homo	2.54				
	111434	R01608	Hs.142736	ESTs				9.80	
	111439	AI476429	Hs.19238	ESTs				10.40	
	111540	U82670	Hs.9786	zinc finger protein 275			15.40		
	111597	R11499	Hs.189716	ESTs				9.20	
40	111895	T80581	Hs.12723	Homo sapiens clone 25153 mRNA sequence		6.80			
	111929	AF027208	Hs.112360	prominin (mouse)-like 1				14.67	
	112054	R43590		gb:yc85g02.s1 Soares infant brain 1NIB H		10.80			
	112210	R49645	Hs.7004	ESTs				10.20	
	112244	AB029000	Hs.70823	KIAA1077 protein	2.99				
45	112382	R59904		gb:yh07g12.s1 Soares infant brain 1NIB H		6.60			
	112392	R60763	Hs.193274	ESTs, Moderately similar to I57588 HSrel			7.10		
	112442	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	3.00				
	112539	R70318	Hs.339730	ESTs				37.20	
	112772	AI992283	Hs.35437	ESTs, Moderately similar to I38026 MLN 6				14.60	
50	112869	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin					4.83
	112935	R71449	Hs.268760	ESTs	2.73				
	112970	AA694010	Hs.6932	Homo sapiens clone 23809 mRNA sequence				12.00	
	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	11.50				
	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f			10.89		
55	113063	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	15.00				
	113073	N39342	Hs.103042	microtubule-associated protein 1B			15.31		
	113078	T40444	Hs.118354	CAT56 protein		7.00			
	113238	R45467	Hs.189813	ESTs				41.20	
	113591	T91881	Hs.200597	KIAA0563 gene product				9.40	
60	113702	T97307		gb:ye53h05.s1 Soares fetal liver spleen	25.00				
	113844	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE				13.91	
	113984	R96696	Hs.35598	ESTs		7.80			
	114073	R44953	Hs.22908	Homo sapiens mRNA; cDNA DKFZp434J1027 (f		7.20			
	114162	AF155661	Hs.22265	pyruvate dehydrogenase phosphatase	3.42				
65	114208	AL049466	Hs.7859	ESTs			6.74		
	114251	H15261	Hs.21948	ESTs				33.20	
	114285	R44338	Hs.22974	ESTs				13.20	
	114313	H18456	Hs.27946	ESTs				10.00	
	114339	AA782845	Hs.22790	ESTs		7.80			
70	114407	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f					4.14
	114560	AI452469	Hs.165221	ESTs				9.80	
	114699	AA127386		gb:zn90d09.r1 Stratagene lung carcinoma		7.60			
	114767	AI859865	Hs.154443	minichromosome maintenance deficient (S	3.21				
	114793	AA158245		gb:zo76c03.s1 Stratagene pancreas (93720			6.00		
75	114833	AI417215	Hs.87159	hypothetical protein FLJ12577				11.40	
	115047	BE270930	Hs.82916	chaperonin containing TCP1, subunit 6A (					4.31
	115060	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3					4.03
	115097	AA256213	Hs.72010	ESTs				35.40	
	115113	AA256460		gb:zr81a04.s1 Soares_NhHMPu_S1 Homo sapi				15.20	
80	115123	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m					4.19
	115134	AW968073	Hs.194331	ESTs, Highly similar to A55713 inositol				12.40	
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	25.00				
	115347	AA356792	Hs.334824	hypothetical protein FLJ14825		7.00			
	115414	AA662240	Hs.283099	AF15q14 protein	3.25				
85	115522	BE614387	Hs.333893	c-Myc target JPO1	3.68				
	115536	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	10.50				
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1				24.40	
	115645	AI207410	Hs.69280	Homo sapiens, clone IMAGE:3636299, mRNA,	4.17				
	115648	AW016811	Hs.234478	Homo sapiens cDNA: FLJ22648 fis, clone H			6.00		



	115652	BE093589	Hs.38178	hypothetical protein FLJ23468	3.81				
	115697	D31382	Hs.63325	transmembrane protease, serine 4	62.14				
	115793	AA424883	Hs.70333	hypothetical protein MGC10753				11.80	
	115816	BE042915	Hs.287588	Homo sapiens cDNA FLJ13675 fis, clone PL				9.71	
5	115892	AA291377	Hs.50831	ESTs			27.40		
	115906	AI767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	2.53				
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	11.82				
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970				34.29	
	115978	AL035864	Hs.69517	cDNA for differentially expressed CO16 g					8.23
10	115985	AA447709	Hs.268115	ESTs, Weakly similar to T08599 probable	3.00				
	116090	AI591147	Hs.61232	ESTs	5.17				
	116096	AA682382	Hs.59982	ESTs			8.20		
	116127	AF126743	Hs.279884	DNAJ domain-containing			10.60		
	116157	BE439838	Hs.44298	mitochondrial ribosomal protein S17					5.82
15	116190	AI949095	Hs.67776	ESTs, Weakly similar to T22341 hypotheti					4.08
	116278	NM_003686	Hs.47504	exonuclease 1	9.50				
	116335	AK001100	Hs.41690	desmocollin 3	3.67				
	116496	AW450694	Hs.21433	hypothetical protein DKFZp547J036			7.00		
	116503	AI925316	Hs.212617	ESTs				12.60	
20	116674	AI768015	Hs.92127	ESTs			32.00		
	116929	AA586922	Hs.80475	polymerase (RNA) II (DNA directed) polyp			7.60		
	116973	AI702054	Hs.166982	phosphatidylinositol glycan, class F			9.80		
	116993	AI417023	Hs.40478	ESTs				10.20	
	117079	H92325		gb:ys85f05.s1 Soares retina N2b4HR Homo				15.20	
25	117317	AI263517	Hs.43322	ESTs				13.40	
	117326	N23629	Hs.241420	Homo sapiens mRNA for KIAA1756 protein,				20.60	
	117396	W20128	Hs.296039	ESTs				10.60	
	117412	N32536	Hs.42645	ESTs				16.00	
	117519	N32528	Hs.146286	kinesin family member 13A				9.11	
30	117693	AW179019	Hs.112110	mitochondrial ribosomal protein L42					4.01
	117721	N46100	Hs.93939	EST				19.80	
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	2.71				
	117903	AA768283	Hs.47111	ESTs				17.80	
	117992	AI015709	Hs.172089	Homo sapiens mRNA; cDNA DKFZp586I2022 (f					4.17
35	118013	AI674126	Hs.94031	ESTs				10.60	
	118017	AI813444	Hs.42197	ESTs			8.82		
	118186	N22886	Hs.42380	ESTs			7.00		
	118325	AI868065	Hs.166184	intersectin 2				13.80	
	118367	N64269	Hs.48946	EST				6.14	
40	118368	N64339	Hs.48956	gap junction protein, beta 6 (connexin 3	3.14				
	118472	AL157545	Hs.42179	bromodomain and PHD finger containing, 3			12.40		
	118709	AA232970	Hs.293774	ESTs				12.20	
	119025	BE003760	Hs.55209	Homo sapiens mRNA; cDNA DKFZp434K0514 (f	4.50				
	119027	AF086161	Hs.114611	hypothetical protein FLJ11808	3.22				
45	119052	R10889		gb:yf38d02.s1 Soares fetal liver spleen			9.60		
	119164	AF221993	Hs.46743	McKusick-Kaufman syndrome			6.60		
	119186	AI979147	Hs.101265	hypothetical protein FLJ22593				10.80	
	119243	T12603		gb:CHR90123 Chromosome 9 exon II Homo sa				9.44	
	119490	AA195276	Hs.263858	ESTs, Moderately similar to B34087 hypot				11.80	
50	119499	AI918906	Hs.55080	ESTs			14.80		
	119599	W45552		gb:zc26d03.s1 Soares_senescent_fibroblas			12.60		
	119780	NM_016625	Hs.191381	hypothetical protein	17.00				
	119845	W79123	Hs.58561	G protein-coupled receptor 87	13.50				
55	119941	AA699485	Hs.58896	ESTs			8.00		
	119994	AA642402	Hs.59142	ESTs	7.73				
	120102	W67353	Hs.170218	KIAA0251 protein			39.60		
	120104	AK000123	Hs.180479	hypothetical protein FLJ20116	2.91				
	120294	AK000059	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par			8.20		
60	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	8.73				
	120599	AA804448	Hs.104463	ESTs			7.00		
	120699	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos				10.00	
	120715	AA292700		gb:zs59a06.s1 NCI_CGAP_GCB1 Homo sapiens			9.40		
	120821	Y19062	Hs.96870	staufer (Drosophila, RNA-binding protein				13.80	
65	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol			9.00		
	120880	AA360240	Hs.97019	EST			15.60		
	120983	AA398209	Hs.97587	EST			27.66		
	121034	AL389951	Hs.271623	nucleoporin 50kD			20.80		
	121121	AA399371	Hs.189095	similar to SALL1 (sal (Drosophila)-like			22.80		
70	121313	AA402713	Hs.97872	ESTs				10.00	
	121369	AW450737	Hs.128791	CGI-09 protein	25.71				
	121376	AA448103	Hs.187958	solute carrier family 6 (neurotransmitte					5.42
	121476	AA412311	Hs.97903	ESTs			8.30		
	121509	AA868939	Hs.97888	ESTs			8.59		
75	121553	AA412488	Hs.48820	TATA box binding protein (TBP)-associat	18.50				
	121753	AK000552	Hs.323518	WD repeat domain 5	7.00				
	121838	AA425680	Hs.98441	ESTs				10.40	
	121857	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	6.00				
	121991	AA430058	Hs.98649	EST				12.20	
	122089	AW016543	Hs.98682	hypothetical protein FKSG32			8.60		
80	122105	AW241685	Hs.98699	ESTs			6.14		
	122163	AA435702	Hs.98829	EST				10.40	
	122318	AA429743		gb:zv60b05.r1 Soares_testis_NHT Homo sap				18.20	
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	13.50				
	122338	AA443311	Hs.98998	ESTs	4.80				
85	122414	AI313473	Hs.99087	ESTs, Weakly similar to S47073 finger pr			8.00		



5	122512	AF053305	Hs.98658	budding uninhibited by benzimidazoles 1	8.80			
	122516	AA449352	Hs.99217	ESTs			9.40	
	122702	AI220089	Hs.99439	ESTs	9.20			
	122852	AI580056	Hs.98992	ESTs			10.40	
	122925	AW268962	Hs.111335	ESTs	6.80			
10	123005	AW369771	Hs.52620	integrin, beta 8		12.60		
	123044	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro			5.35	
	123160	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypotheti		6.06		
	123315	AA496369		gb:zv37d10.s1 Soares ovary tumor NbHOT H		12.40		
	123329	Z47542	Hs.179312	small nuclear RNA activating complex, po		11.80		
15	123497	AA765256	Hs.135191	ESTs, Weakly similar to unnamed protein	12.00			
	123518	AL035414	Hs.21068	hypothetical protein		13.00		
	123519	AW015887	Hs.112574	ESTs	12.20			
	123614	AK000492	Hs.98806	hypothetical protein		7.80		
	123616	AA680003	Hs.109363	Homo sapiens cDNA: FLJ23603 fis, clone L			10.60	
20	123673	BE550112	Hs.158549	ESTs, Weakly similar to T2D3_HUMAN TRANS	23.00			
	123727	AI083986	Hs.282977	hypothetical protein FLJ13490	7.00			
	123731	AA609839		gb:ae62f01.s1 Stratagene lung carcinoma		9.80		
	123752	AA227714	Hs.179703	KIAA0129 gene product	3.50			
	123900	AA621223	Hs.112953	EST			12.80	
25	124006	AI147155	Hs.270016	ESTs	97.00			
	124059	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti	3.02			
	124069	AF134160	Hs.7327	claudin 1		27.80		
	124191	T96509	Hs.248549	ESTs, Moderately similar to S65657 alpha			35.80	
	124273	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	7.20			
30	124297	AL080215	Hs.102301	Homo sapiens mRNA; cDNA DKFZp586J0323 (f			11.00	
	124305	AW963221		gb:EST375294 MAGE resequences, MAGH Homo			16.00	
	124676	AI360119.comp	Hs.181013	phosphoglycerate mutase 1 (brain)				6.08
	124874	BE550182	Hs.127826	RalGEF-like protein 3, mouse homolog			21.00	
	124904	AK000483	Hs.93872	KIAA1682 protein	9.40			
35	124969	AI650360	Hs.100256	ESTs			10.80	
	125000	T58615	Hs.110640	ESTs			9.80	
	125201	AA693960	Hs.103158	ESTs, Weakly similar to T33296 hypotheti	7.60			
	125266	W90022	Hs.186809	ESTs, Highly similar to LCT2_HUMAN LEUKO	6.59			
	125299	T32982	Hs.102720	ESTs			9.57	
40	125356	AI057052	Hs.133554	ESTs, Weakly similar to Z195_HUMAN ZINC			14.00	
	125370	AA256743	Hs.134158	Homo sapiens, Similar to KIAA0092 gene p		8.20		
	125418	AA777690	Hs.188501	ESTs			13.20	
	125433	AL162066	Hs.54320	hypothetical protein DKFZp762D096	21.40			
	125437	AI609449	Hs.140197	ESTs	6.96			
45	125446	BE219987	Hs.166982	phosphatidylinositol glycan, class F	8.80			
	125711	AA305800	Hs.5672	hypothetical protein AF140225			11.20	
	125756	BE174587	Hs.289721	growth arrest specific transcript 5				4.31
	125757	AI274906	Hs.166835	ESTs, Highly similar to 1814460A p53-ass			15.60	
	125769	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	3.20			
50	125839	AW836261	Hs.337717	ESTs	8.20			
	125850	W85858	Hs.99804	ESTs	2.65			
	125875	H14480		gb:ym18b09.r1 Soares infant brain 1NIB H	7.40			
	125924	BE272506	Hs.82109	syndecan 1				4.23
	125972	AI927475	Hs.35406	ESTs, Highly similar to unnamed protein				3.98
55	126034	H60340		gb:yr39b04.r1 Soares fetal liver spleen			10.60	
	126327	AA432266	Hs.44648	ESTs	11.60			
	126345	N49713		gb:yv23f06.s1 Soares fetal liver spleen	6.67			
	126435	AW614529	Hs.285847	CGI-19 protein			10.60	
	126487	AA283809	Hs.184601	solute carrier family 7 (cationic amino				4.38
60	126521	AI475110	Hs.203933	ESTs	6.60			
	126522	W31912		gb:zc76d03.s1 Pancreatic Islet Homo sapi			14.80	
	126543	AL035864	Hs.69517	cDNA for differentially expressed CO16 g				4.01
	126567	AA058394	Hs.57887	ESTs, Weakly similar to KIAA0758 protein		7.80		
	126605	AA676910		gb:zj65h07.s1 Soares_fetal_liver_spleen_			11.60	
65	126627	AA497044	Hs.20887	hypothetical protein FLJ10392			14.60	
	126628	N49776	Hs.170994	hypothetical protein MGC10946	8.00			
	126737	AW976516	Hs.283707	Homo sapiens cDNA: FLJ21354 fis, clone C	2.92			
	126795	AW975076	Hs.172589	nuclear phosphoprotein similar to S. cer	7.50			
	126802	AW805510	Hs.97056	hypothetical protein FLJ21634	11.60			
70	126892	AF121856	Hs.284291	sorting nexin 6	3.50			
	126928	AA480902	Hs.137401	ESTs			22.83	
	126979	AA210954		gb:zq89h10.r1 Stratagene hNT neuron (937			11.80	
	126986	AI279892	Hs.46801	sorting nexin 14			11.60	
	126992	AI809521		gb:wf30e03.x1 Soares_NFL_T_GBC_S1 Homo s			20.80	
75	127066	R25066		gb:yq42c07.r1 Soares infant brain 1NIB H			27.60	
	127099	AA347668		gb:EST54026 Fetal heart II Homo sapiens			21.60	
	127139	AA830233	Hs.293585	ESTs			11.20	
	127209	AA305023	Hs.81964	SEC24 (S. cerevisiae) related gene famii	3.10			
	127221	BE062109	Hs.241551	chloride channel, calcium activated, fam	2.76			
80	127225	AA315933	Hs.120879	ESTs			16.80	
	127313	AK002014	Hs.47546	Homo sapiens cDNA FLJ11458 fis, clone HE	14.00			
	127444	AW978474	Hs.7560	Homo sapiens mRNA for KIAA1729 protein,			13.60	
	127500	AW971353	Hs.162115	ESTs	11.20			
	127524	AI243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin		7.80		
85	127540	N45572	Hs.105362	Homo sapiens, clone MGC:18257, mRNA, com	3.53			
	127599	AA613204	Hs.150399	ESTs			13.80	
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture			28.00	
	127662	W80755	Hs.8294	KIAA0196 gene product			19.80	
	127668	AI343257	Hs.139993	ESTs			11.20	



	127746	AI239495	Hs.120189	ESTs			14.18	
	127812	AA741368	Hs.291434	ESTs	4.50			
	127817	AA836641	Hs.163085	ESTs			24.60	
5	127959	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L			9.20	
	127960	AI613226	Hs.41569	phosphatidic acid phosphatase type 2A			16.83	
	127969	F06498	Hs.93748	Homo sapiens cDNA FLJ14676 fis, clone NT	13.60			
	128015	Z21169	Hs.334659	hypothetical protein MGC14139	7.00			
	128027	AI433721	Hs.164153	ESTs			37.40	
	128077	AI310330	Hs.128720	ESTs			9.60	
10	128166	NM_006147	Hs.11801	interferon regulatory factor 6			9.24	
	128226	AI284940	Hs.289082	GM2 ganglioside activator protein	19.00			
	128305	AI954968	Hs.279009	matrix Gla protein			10.40	
	128341	AA191420	Hs.185030	ESTs	9.00			
15	128527	AA504583	Hs.101047	transcription factor 3 (E2A immunoglobul			4.30	
	128539	R46163	Hs.258618	ESTs	12.60			
	128568	H12912	Hs.274691	adenylate kinase 3			4.56	
	128572	AA933022	Hs.256583	interleukin enhancer binding factor 3, 9			10.00	
	128777	AI878918	Hs.10526	cysteine and glycine-rich protein 2		16.80		
20	128781	N71826	Hs.105465	small nuclear ribonucleoprotein polypept			4.48	
	128796	AJ000152	Hs.105924	defensin, beta 2	8.12			
	128920	AA622037	Hs.166468	programmed cell death 5			4.62	
	128924	BE279383	Hs.26557	plakophilin 3			4.04	
	128971	H05132	Hs.107510	ESTs	12.60			
25	129008	AL079648	Hs.301088	ESTs	8.80			
	129041	BE382756	Hs.169902	solute carrier family 2 (facilitated glu			6.05	
	129075	BE250162	Hs.83765	dihydrofolate reductase	2.59			
	129105	AI769160	Hs.108681	Homo sapiens brain tumor associated prot		6.67		
	129189	AB023179	Hs.9059	KIAA0962 protein	8.00			
30	129229	AF013758	Hs.109643	polyadenylate binding protein-interactin	4.00			
	129241	AI878857	Hs.109706	hematological and neurological expressed			4.06	
	129300	W94197	Hs.110165	ribosomal protein L26 homolog	2.55			
	129404	AI267700	Hs.317584	ESTs	18.00			
	129457	X61959	Hs.207776	aspartylglucosaminidase	6.50			
35	129466	L42583	Hs.334309	keratin 6A	12.94			
	129494	AI148976	Hs.112062	ESTs			11.00	
	129605	AF061812	Hs.115947	keratin 16 (focal non-epidermolytic palm			4.46	
	129641	AI911527	Hs.11805	ESTs			12.00	
	129665	AW163331	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic retic			4.70	
40	129703	BE388665	Hs.179999	Homo sapiens, clone IMAGE:3457003, mRNA			4.02	
	129720	AA156214	Hs.12152	APMCF1 protein			5.71	
	129748	M16707	Hs.123053	H4 histone, family 2	3.50			
	129890	AI868872	Hs.282804	hypothetical protein FLJ22704			4.21	
	129896	BE295568	Hs.13225	UDP-Gal:betaGlcNAc beta 1,4- galactosylt	2.56			
45	129945	BE514376	Hs.165998	PAI-1 mRNA-binding protein			4.03	
	130010	AA301116	Hs.142838	nucleolar phosphoprotein Nopp34		7.00		
	130026	T40480	Hs.332112	EST	6.40			
	130080	X14850	Hs.147097	H2A histone family, member X			4.65	
50	130149	AW067805	Hs.172665	methylenetetrahydrofolate dehydrogenase	2.74			
	130285	AA063546	Hs.75981	ubiquitin specific protease 14 (tRNA-gua		7.40		
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic			3.91	
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	4.87			
	130500	AB007913	Hs.158291	KIAA0444 protein			9.60	
55	130524	U89995	Hs.159234	forkhead box E1 (thyroid transcription f		13.40		
	130541	X05608	Hs.211584	neurofilament, light polypeptide (68kD)		8.20		
	130553	AF062649	Hs.252587	pituitary tumor-transforming 1			6.06	
	130567	AA383092	Hs.1608	replication protein A3 (14kD)		7.00		
	130577	M69241	Hs.162	insulin-like growth factor binding prote	3.04			
60	130627	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	3.87			
	130648	AI458165	Hs.17296	hypothetical protein MGC2376			16.20	
	130697	L29472	Hs.1802	major histocompatibility complex, class			17.80	
	130744	H59696	Hs.18747	POP7 (processing of precursor, S. cerevi			5.28	
	130800	AI187292	Hs.19574	hypothetical protein MGC5469			4.43	
	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	16.84			
65	130869	J03626	Hs.2057	uridine monophosphate synthetase (orotat			4.92	
	130925	AF093419	Hs.169378	multiple PDZ domain protein			9.60	
	130994	W17044	Hs.327337	ESTs	12.40			
	131028	AI879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	10.21			
	131031	NM_001650	Hs.288650	aquaporin 4			9.80	
70	131041	T15767	Hs.22452	Homo sapiens mRNA for KIAA1737 protein,			9.60	
	131058	W28545	Hs.101514	hypothetical protein FLJ10342			17.00	
	131090	AI143139	Hs.2288	visinin-like 1	2.74			
	131112	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f		8.80		
	131148	AW953575	Hs.303125	p53-induced protein PIGPC1	3.12			
75	131185	BE280074	Hs.23960	cyclin B1	3.07			
	131200	BE540516	Hs.293732	hypothetical protein MGC3195	3.07			
	131219	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	2.87			
	131257	AW339037	Hs.24908	ESTs			14.67	
	131375	AW293165	Hs.143134	ESTs		19.20		
80	131460	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.50			
	131476	AI521663	Hs.334644	hypothetical protein FLJ14668	15.00			
	131510	BE245374	Hs.27842	hypothetical protein FLJ11210		7.80		
	131646	BE302464	Hs.30057	MRS2 (S. cerevisiae)-like, magnesium hom		7.00		
	131786	BE000971	Hs.306083	Novel human gene mapping to chromosome 22	2.65			
85	131839	AB014533	Hs.33010	KIAA0633 protein			35.20	
	131843	AA192315	Hs.184062	putative Rab5-interacting protein			4.11	



	131877	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	19.00			
	131885	BE502341	Hs.3402	ESTs	6.48			
	131921	AA456093	Hs.34720	ESTs		8.40		
5	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	56.00			
	131958	NM_014062	Hs.3566	ART-4 protein				3.82
	131965	W79283	Hs.35962	ESTs	3.03			
	132000	AW247017	Hs.36978	melanoma antigen, family A, 3		9.80		
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	3.30			
10	132109	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	21.00			
	132114	NM_006152	Hs.40202	lymphoid-restricted membrane protein		8.40		
	132162	AA315805	Hs.94560	desmoglein 2				12.25
	132164	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio	2.70			
	132180	NM_004460	Hs.418	fibroblast activation protein, alpha	2.71			
15	132181	AW961231	Hs.16773	Homo sapiens clone TCCCA00427 mRNA sequ	3.83			
	132182	NM_014210	Hs.70499	ecotropic viral integration site 2A				13.20
	132231	AA662910	Hs.42635	hypothetical protein DKFZp434K2435	9.50			
	132277	AK001745	Hs.184628	hypothetical protein FLJ10883	4.50			
	132328	NM_014787	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe				9.20
20	132394	AK001680	Hs.30488	DKFZP434F091 protein				19.80
	132424	AA417878	Hs.48401	ESTs, Moderately similar to ALU8_HUMAN A		8.60		
	132528	T78736	Hs.50758	SMC4 (structural maintenance of chromoso		27.40		
	132543	BE568452	Hs.5101	protein regulator of cytokinesis 1	4.38			
	132544	L19778	Hs.51011	H2A histone family, member P		7.00		
25	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	2.64			
	132552	BE621985	Hs.296922	thiopurine S-methyltransferase				15.83
	132581	AK000631	Hs.52256	hypothetical protein FLJ20624		6.60		
	132617	AF037335	Hs.5338	carbonic anhydrase XII	4.95			
	132638	AI796870	Hs.54277	DNA segment on chromosome X (unique) 992		8.20		
30	132653	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	4.38			
	132669	W38586	Hs.293981	guanine nucleotide binding protein (G pr				4.36
	132710	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	4.60			
	132771	Y10275	Hs.56407	phosphoserine phosphatase	3.71			
	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,				9.48
35	132833	U78525	Hs.57783	eukaryotic translation initiation factor				5.83
	132892	AW834050	Hs.9973	tensin				12.00
	132906	BE613337	Hs.234896	geminin	3.09			
	132959	AW014195	Hs.61472	ESTs, Weakly similar to YAE6_YEAST HYPOT				3.87
	132962	AA576635	Hs.6153	CGI-48 protein	3.50			
40	132990	X77343	Hs.334334	transcription factor AP-2 alpha (activat	6.18			
	132994	AA112748	Hs.279905	clone HQ0310 PRO0310p1	3.19			
	133000	AL042444	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	2.96			
	133050	X73424	Hs.63788	propionyl Coenzyme A carboxylase, beta p	2.55			
	133083	BE244588	Hs.6456	chaperonin containing TCP1, subunit 2 (b				4.00
45	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso				8.96
	133134	AF198620	Hs.65648	RNA binding motif protein 8A				4.28
	133155	M58583	Hs.662	cerebellin 1 precursor				10.80
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	3.00			
	133204	BE267696	Hs.254105	enolase 1, (alpha)				4.63
50	133412	U41493	Hs.73112	guanine nucleotide binding protein (G pr		12.50		
	133421	AF134160	Hs.7327	claudin 1	2.85			
	133451	AW970026	Hs.73818	ubiquinol-cytochrome c reductase hinge p				4.66
	133453	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept		6.80		
	133504	NM_004415	Hs.74316	desmoplakin (DPI, DPII)	6.14			
55	133506	BE562958	Hs.74346	hypothetical protein MGC14353				4.55
	133615	M62843	Hs.75236	ELAV (embryonic lethal, abnormal vision,				17.80
	133627	NM_002047	Hs.75280	glycyl-tRNA synthetase				4.85
	133649	U25849	Hs.75393	acid phosphatase 1, soluble				6.34
	133669	NM_006925	Hs.166975	splicing factor, arginine/serine-rich 5				14.00
60	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		6.11		
	133776	BE268649	Hs.177766	ADP-ribosyltransferase (NAD+; poly (ADP-				4.91
	133865	AB011155	Hs.170290	discs, large (Drosophila) homolog 5	3.07			
	133946	AJ001258	Hs.173878	NIPSNAP, C. elegans, homolog 1				4.60
	133973	N55540	Hs.78026	ESTs, Weakly similar to similar to ankyr				13.00
65	134047	BE262529	Hs.78771	phosphoglycerate kinase 1				3.85
	134098	BE513171	Hs.79086	mitochondrial ribosomal protein L3	2.56			
	134107	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte		8.20		
	134112	AW449809	Hs.79150	chaperonin containing TCP1, subunit 4 (d				4.08
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	31.00			
70	134160	T98152	Hs.79432	fibrillin 2 (congenital contractual ara		24.60		
	134168	AA398908	Hs.181634	Homo sapiens cDNA: FLJ23602 fis, clone L				6.71
	134185	AA285136	Hs.301914	neuronal specific transcription factor D				14.74
	134201	L35035	Hs.79886	ribose 5-phosphate isomerase A (ribose 5		8.40		
	134272	X76040	Hs.278614	protease, serine, 15	4.50			
75	134276	BE083936	Hs.80976	antigen identified by monoclonal antibod		9.00		
	134353	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m				16.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	2.80			
	134380	AU077143	Hs.179565	minichromosome maintenance deficient (S.	4.68			
	134423	H53497	Hs.83006	CGI-139 protein				3.84
80	134469	AA279661	Hs.83753	small nuclear ribonucleoprotein polypept				5.81
	134470	X54942	Hs.83758	CDC28 protein kinase 2				4.21
	134498	AW246273	Hs.84131	threonyl-tRNA synthetase				7.30
	134502	BE148534	Hs.84168	UV-B repressed sequence, HUR 7		13.60		
	134510	NM_002757	Hs.250870	mitogen-activated protein kinase				9.70
	134548	N95406	Hs.333495	Deleted in split-hand/split-foot 1 regio				4.63
85	134654	AK001741	Hs.8739	hypothetical protein FLJ10879	6.00			



	134724	AF045239	Hs.321576	ring finger protein 22		12.00
	134743	AA044163	Hs.89463	potassium large conductance calcium-acti	4.00	
	134781	AA374372	Hs.89626	parathyroid hormone-like hormone		25.20
5	134806	AD001528	Hs.89718	spermine synthase		4.58
	134853	BE268326	Hs.90280	5-aminoimidazole-4-carboxamide ribonucle		4.79
	134859	D26488	Hs.90315	KIAA0007 protein		6.20
	134891	R51083	Hs.90787	ESTs		7.40
	134960	BE246400	Hs.285176	acetyl-Coenzyme A transporter	4.00	
	134993	BE409809	Hs.301005	purine-rich element binding protein B		4.48
10	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	9.50	
	135080	AI761180	Hs.94211	rcd1 (required for cell differentiation,	5.00	
	135103	NM_003428	Hs.9450	zinc finger protein 84 (HPF2)		11.00
	135145	AW014729	Hs.95262	nuclear factor related to kappa B bindin		4.01
	135184	U13222	Hs.96028	forkhead box D1		7.00
15	135242	AI583187	Hs.9700	cyclin E1	13.50	
	135286	AW023482	Hs.97849	ESTs	6.46	
	135289	AW372569	Hs.9788	hypothetical protein MGC10924 similar to		8.80
	135355	AK001652	Hs.99423	ATP-dependent RNA helicase	10.00	
	135371	NM_006025	Hs.997	protease, serine, 22	8.00	
20	135393	L11244	Hs.99886	complement component 4-binding protein,		14.60

TABLE 5B shows the accession numbers for those primekeys lacking unigenelD's for Table 5A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
35	117079	1621717_1 H92325 T97125
	124305	242183_1 AW963221 AA344870 AA344871 H93331
	101502	18202_6 M26958
	109792	754958_1 R49625 F10674
	126034	1598157_1 H60340 N91637
40	102768	44641_1 U82321 H66077
	126345	1653833_1 N49713 N49819 W03810
	127066	1703458_1 R25066 R20144 R20145 Z43845
	127099	244301_1 AA347668 AW956810 Z44271 F07065 F07064 R13506
	119243	1774795_1 T12603 T12604
45	125875	1566433_1 H14480 N98295
	112054	1538292_1 R43590 F10439
	126979	171411_1 AA210954 AA211007
	126992	880655_1 AI809521 H12174 Z42556
	122318	292419_1 AA429743 AA442754
	114699	135322_1 AA127386 R15644 AA127404
50	114793	150742_1 AA158245 AA158235
	108305	111550_1 AA071391 AA069892 AA069891
	108393	113411_1 AA075211 AA075245 AA075126 AA074946
	100867	tigr_HT4586 U14622
55	123731	genbank_AA609839 AA609839
	109700	genbank_F09609 F09609
	120715	genbank_AA292700 AA292700
	113702	genbank_T97307 T97307
	115113	genbank_AA256460 AA256460
60	101045	entrez_J05614 J05614
	108554	genbank_AA084948 AA084948
	108573	genbank_AA086005 AA086005
	119052	149538_1 R10889 R10888
	126522	416020_1 W31912 AI167491
65	126605	439280_1 AA676910 AA778853 AA778865 W86800
	103768	46922_1 W42667 AI580740 AI690440 AI561350 AW467906 AW151450 AI825927 AL041716 AI885600 AI742213 AW248624 AI955498 AA033947
		AA845593 AI623711 N68583 C00064 AA193567 AW083868 AW163216 AA191595 AA522778 AI628008 AI915518 AA843508 AI926195
		AA176265 AW167963 AA992115 W93647 AW103572 AI862994 AI342059 AA911719 AA176155 AA024712 AA069988 AA205591 AI591107
		AI199673 AI811766 AI275832 AI422233 AI191852 AI096682 AI580124 AI683612 AA582453 AA927559 AA486415 T32414 AI084978 H44849
70		H44848 H20477 T91695 W47039 AA070055 AA024795 AA328855 AA379248 AA379330 AA385580 W25920 W03688 AA448359 AA093881
		AW362477 AA089997 AI350265 W93479 N99688 AA932257 AW351469 H68590 AA663402 AA069771 AW087986 AI858420 AA600214
		AI970774 AI857712 AI683081 AI885584 AW131150 AI567981 AW002714 AW189973 AW075495 AW168303 AA953714 AW516881 AI357375
		AI566663 AW512676 AI570580 AI023690 AA448216 AI079853 AI422707 AA779516 AW026972 AW130082 AW162307 AW438646 AA709332
		AW192394 AI167350 AI217879 AI129152 AA719509 AI350480 AA663418 AI003634 AW118546 AA180261 AA442833 AI268625 AA888881
75		AI038759 AA846723 AI248770 AA993694 AI280335 AI885107 AW518649 AA641563 AA995835 AA582521 AI276744 AA436478 AI017360
		AI620763 AI859887 N73926 AI076327 AI741615 AI160617 AW172819 AI492005 AA677429 AA996334 AI693771 AI950039 AI245629 AI288515
		AI866186 T93293 AA173262 AA599779 AI680092 AW439316 AI084555 AI272672 AI583507 AW473219 AA738132 AW473283 AI367492
		AA995410 AI689624 AA206353 AI033095 AI040382 AA873630 AI221074 AI934840 AI418680 AA844306 R94503 AA773520 AA843169
		AA219425 AA629658 AI811719 AW411275 AI590981 W37907 AI591178 AI684051 AA983238 AA669347 AA976239 AA704570 AI628339
80		AI884391 AI241580 AI003539 AW176687 AA009650 N34566 AI333493 AI186070 AA070827 AA411683 AI280884 AA872023 AA207255
		AA021576 N71953 AI885888 AW076039 T15777 AI537673 AW248048 H09554 W93480 W47001 AW079114 AA063160 AA757453 R60788
		AI859431 H20478 AA218882 AA757465 AA100995 AI864135 AI934209 AA070503 H47008 AA219646 W61039 W93907 AW385050 W37967
		W78028 AA189007 AA479136 R93650 AA442312 T30287 AA847628 AA180262 AA009649 C03892 AW149464 AA310963 AA219693
		AA069747 R29207 AA094784 AA293615 AA447848 AI984167 N90393 C05097 N56499 AW292351 AW149681 AW473258 AA629322 AI004409
85		AW105577 AI954937 AI811070 AA902422 AW514437 AA535460 AA916877 AW517122 AA974657 AA975649 AW517130 AW517129 F31737
		W07688 AA193645 AA378994 AA489273 F32267 W39303 AA021181 N86810 AA406524 AA062553 AA436801 H08985 H15979 N40310



AA436789 AA232172 AW360778 W25862 R60282 AA436530 AA378894 AA187461 AI940535 AA604210 AA089514 AA360421 N88243 N84281  
AA209340 N56174 N88374 AA191088 AW247691 AA249013 AA093111 AA972536 AW298594 AA375893 T12139 W28186 AW243849  
AI288629 AA843996 W15260 AI188286 AW248079 R15836

119599 genbank\_W45552 W45552  
112382 genbank\_R59904 R59904  
105264 genbank\_AA227934 AA227934  
100071 entrez\_A28102 A28102  
123315 714071\_1 AA496369 AA496646

Table 6A shows 99 genes up-regulated nonsmokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

Pkey: Unique Eos probeset identifier number  
ExAccn: Exemplar Accession number, Genbank accession number  
UnigeneID: Unigene number  
Unigene Title: Unigene gene title  
R1: average of AI for samples from non-smokers with adenocarcinoma divided by the 90th percentile of AI for samples from smokers with adenocarcinoma  
R2: average of AI for samples from non-smokers with squamous cell carcinoma divided by the 90th percentile of AI for samples from smokers with squamous cell carcinoma

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		3.64
101174	L17330	Hs.280	pre-T/NK cell associated protein	15.00	
101296	Y12490	Hs.85092	thyroid hormone receptor interactor 11		2.46
101304	AA001021	Hs.6685	thyroid hormone receptor interactor 8		12.00
101806	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias		2.68
101972	S82472		gb:beta -pol=DNA polymerase beta (exon a		2.11
102274	U30930	Hs.158540	UDP glycosyltransferase 8 (UDP-galactose	7.50	
102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	7.50	
102832	U92015		gb:Human clone 143789 defective mariner	13.50	
103010	X52509	Hs.161640	tyrosine aminotransferase	9.50	
103439	X98266		gb:H.sapiens mRNA for ligase like protei		2.50
103563	L02911	Hs.150402	activin A receptor, type I	9.00	
103857	AI076795	Hs.45033	lacrimal proline rich protein		3.94
104239	AB002367	Hs.21355	doublecortin and CaM kinase-like 1	13.50	
104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m		12.66
104907	AA055829	Hs.196701	ESTs, Weakly similar to ALU1_HUMAN ALU	16.50	
106131	BE514788	Hs.296244	SNARE protein		2.17
106672	H47233	Hs.30643	ESTs	7.00	
106872	T56887	Hs.18282	KIAA1134 protein	11.50	
106960	AA156238	Hs.32501	ESTs		2.38
106971	Z43846	Hs.194478	Homo sapiens mRNA; cDNA DKFZp434O1572 (f	9.50	
107982	AA035375	Hs.57887	ESTs, Weakly similar to KIAA0758 protei		2.95
108562	AA100796		gb:zm26c06.s1 Stratagene pancreas (93720	16.50	
108599	AB018549	Hs.69328	MD-2 protein	13.00	
108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypotheti		2.40
109247	AA314907	Hs.85950	ESTs	7.00	
109630	R44607	Hs.22672	ESTs		5.00
110193	AI004874	Hs.310764	Homo sapiens mRNA; cDNA DKFZp434M082 (fr	12.50	
110234	H24458	Hs.32085	EST	16.50	
110644	R94207	Hs.268989	ESTs, Highly similar to type II CALM/AF1	8.00	
110886	AW274992	Hs.72249	three-PDZ containing protein similar to	17.00	
111057	T79639	Hs.14629	ESTs	16.50	
111950	AF071594	Hs.110457	Wolf-Hirschhorn syndrome candidate 1	11.00	
112291	R53972	Hs.26026	ESTs		3.00
112956	Z43784	Hs.75893	ankyrin 3, node of Ranvier (ankyrin G)		2.79
113009	T23699	Hs.7246	ESTs		4.50
113060	BE564162	Hs.250820	hypothetical protein FLJ14827	9.79	
113073	N39342	Hs.103042	microtubule-associated protein 1B	32.50	
113074	AK001335	Hs.31137	protein tyrosine phosphatase, receptor t		3.82
113121	T48011	Hs.8764	EST		2.21
113125	AA968672	Hs.8929	hypothetical protein FLJ11362	19.50	
113757	AA703095	Hs.18631	ESTs		2.65
113848	W52854	Hs.27099	hypothetical protein FLJ23293 similar to	6.00	
113884	AI333076	Hs.28529	chromosome 12 open reading frame 2		6.00
113936	W17056	Hs.83623	nuclear receptor subfamily 1, group I, m		4.63
114875	AA235609	Hs.236443	Homo sapiens mRNA; cDNA DKFZp564N1063 (		7.00
114987	AA251016	Hs.87808	EST		6.00
115460	AW958439	Hs.38613	ESTs		2.27
115722	W91892	Hs.59609	ESTs		9.00
116261	AA481788	Hs.190150	ESTs	9.50	
116830	H61037	Hs.70404	ESTs, Weakly similar to ALU2_HUMAN ALU	8.50	
116970	AB023179	Hs.9059	KIAA0962 protein	7.50	
117178	H98675	Hs.269034	ESTs		2.68
117757	AF088019	Hs.46732	EST	7.50	
118283	AA287747	Hs.173012	ESTs, Weakly similar to A46010 X-linked	16.50	
118384	AF217525	Hs.49002	Down syndrome cell adhesion molecule		2.50
118657	AI822106	Hs.49902	ESTs		2.39
120328	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapi		3.50
120404	AB023230	Hs.96427	KIAA1013 protein	7.00	
120524	AA261852	Hs.192905	ESTs	6.00	
120688	AW207555	Hs.97093	Homo sapiens cDNA: FLJ23004 fis, clone L	17.92	



	121558	AA412497	gb:zt95g12.s1 Soares_testis_NHT Homo sap		2.95
	121676	H56037	ESTs	10.00	
	121936	AI024600	ESTs	15.00	
5	121938	AA428659	ESTs	14.00	
	122177	AA435789	EST	8.93	
	123442	AA299652	Homo sapiens cDNA FLJ11643 fis, clone HE	13.04	
	123551	AA608837	gb:af03h12.s1 Soares_testis_NHT Homo sap	11.50	
	123756	AA609971	EST	11.00	
10	123861	AA620840	gb:af89g01.s1 Soares_testis_NHT Homo sap		2.50
	124371	N24924	ESTs	6.50	
	127477	BE328720	ESTs		4.33
	127591	AI190540	ESTs		3.02
	128252	AA455924	ESTs	7.00	
15	128426	AI265784	ESTs		2.08
	128925	R67419	Homo sapiens cDNA FLJ12900 fis, clone NT		2.11
	128945	AI990506	Homo sapiens mRNA; cDNA DKFZp547E184 (fr	10.00	
	129105	AI769160	Homo sapiens brain tumor associated prot	15.50	
	129235	AW977238	KIAA1055 protein		4.25
20	129506	AB020684	KIAA0877 protein	6.50	
	129595	U09550	oviductal glycoprotein 1, 120kD (mucin 9		10.00
	130160	AA305688	UDP-Gal:betaGlcNAc beta 1,3-galactosyltr	20.00	
	130340	D82326	solute carrier family 3 (cystine, dibasi	11.50	
	131220	AB023194	KIAA0977 protein	17.50	
	131430	AI879148	fatty acid binding protein 7, brain	6.10	
25	132114	NM_006152	lymphoid-restricted membrane protein		6.15
	132458	AA935315	Homo sapiens cDNA: FLJ21693 fis, clone C		5.58
	132647	NM_006927	sialyltransferase 4B (beta-galactosidase	7.50	
	132655	D49372	small inducible cytokine subfamily A (Cy		2.53
30	132682	AI077500	serologically defined colon cancer antig		2.50
	132747	AA345241	ESTs, Weakly similar to KIAA1330 protein		2.83
	132812	R50333	Leman coiled-coil protein		3.82
	133337	AF085983	ESTs		5.00
	133876	AL134906	phosphorylase, glycogen; liver (Hers dis		3.00
35	134119	AW157837	fasciculation and elongation protein zet		2.06
	134464	AA302983	CCR4-NOT transcription complex, subunit		2.27
	134542	M14156	insulin-like growth factor 1 (somatomedi		11.50
	135002	AA448542	G antigen 7B	87.00	
	135305	AA203555	Homo sapiens cDNA FLJ14903 fis, clone PL		6.50

TABLE 6B show the accession numbers for those primekeys lacking unigenelD's for Table 6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

	Pkey	CAT number	Accessions
55	108562	36375_1	AA100796 AF020589 AA074629 AA075946 AA100849 AA085347 AA126309 AA079311 AA079323 AA085274
	103439	35330_1	X98266 N41124
	123551	genbank_AA608837	AA608837
	123861	genbank_AA620840	AA620840
	102832	entrez_U92015	U92015
60	101972	entrez_S82472	S82472
	121558	genbank_AA412497	AA412497



Table 7A shows 98 genes down-regulated in non-smokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number				
	ExAccn:	Exemplar Accession number, Genbank accession number				
	UnigeneID:	Unigene number				
	Unigene Title:	Unigene gene title				
10	R1:	90th percentile of AI for samples from smokers with adenocarcinoma divided by the average of AI for samples from non-smokers with adenocarcinoma.				
	R2:	90th percentile of AI for samples from smokers with squamous cell carcinoma divided by the average of AI for samples from non-smokers with squamous cell carcinoma.				
	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
15	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3		164.10
	100380	D82343	Hs.18551	neuroblastoma (nerve tissue) protein		77.40
	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid	102.40	
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte	463.80	
	101046	K01160		(NONE)	672.00	
20	101066	AW970254	Hs.889	Charot-Leyden crystal protein	66.00	
	101175	U82671	Hs.36980	melanoma antigen, family A, 2		77.20
	101497	W05150	Hs.37034	homeo box A5	62.80	
	101663	NM_003528	Hs.2178	H2B histone family, member Q	78.00	
	101677	NM_000715	Hs.1012	complement component 4-binding protein,	186.20	
25	101745	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	80.08	
	101941	S77583		gb:HERVK10/HUMMTV reverse transcriptase	99.20	
	102125	NM_006456	Hs.288215	sialyltransferase		103.10
	102242	U27185	Hs.82547	retinoic acid receptor responder (tazaro	67.00	
	102340	U37055	Hs.278657	macrophage stimulating 1 (hepatocyte gro	71.60	
30	102369	U39840	Hs.299867	hepatocyte nuclear factor 3, alpha		69.70
	102457	NM_001394	Hs.2359	dual specificity phosphatase 4	153.00	
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2		65.70
	102796	AL079646	Hs.107019	symplesin; Huntingtin interacting protei		58.80
	102829	NM_006183	Hs.80962	neurotensin		268.80
35	103207	X72790		gb:Human endogenous retrovirus mRNA for	70.00	
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o		212.10
	103260	X78416	Hs.3155	casein, alpha		130.70
	103351	X89211		gb:H.sapiens DNA for endogenous retrovir	64.60	
40	104212	AB002298	Hs.173035	KIAA0300 protein	66.80	
	104252	AF002246	Hs.210863	cell adhesion molecule with homology to	63.80	
	104258	AF007216	Hs.5462	solute carrier family 4, sodium bicarbon	94.40	
	105024	AA126311	Hs.9879	ESTs	68.20	
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S		74.60
45	106440	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub		71.10
	106566	BE298210		gb:601118016F1 NIH_MGC_17 Homo sapiens c	73.20	
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	83.80	
	106614	AA648459	Hs.335951	hypothetical protein AF301222		62.30
	106654	AW075485	Hs.286049	phosphoserine aminotransferase		202.40
	106999	H93281	Hs.10710	hypothetical protein FLJ20417		89.60
50	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c		66.40
	108810	AW295647	Hs.71331	hypothetical protein MGC5350		95.50
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act		63.40
	109597	AA989362	Hs.293780	ESTs	85.00	
	109691	T65568	Hs.12860	ESTs		58.70
55	109704	AI743880	Hs.12876	ESTs		60.60
	110942	R63503	Hs.28419	ESTs	76.40	
	111722	R23924	Hs.23596	EST	74.60	
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-li	64.80	
60	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f		76.70
	113073	N39342	Hs.103042	microtubule-associated protein 1B		120.20
	114251	H15261	Hs.21948	ESTs	127.20	
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	174.00	
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461		91.00
65	115815	AW905328	Hs.180842	ribosomal protein L13	66.40	
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH		226.60
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970	82.80	
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093		361.60
	116552	D20508	Hs.164649	hypothetical protein DKFZp434H247	69.00	
	116571	D45652		gb:HUMGS02848 Human adult lung 3' direct	64.20	
70	118466	N66741		gb:yz33g08.s1 Morton Fetal Cochlea Homo		63.50
	120484	AA253170	Hs.96473	EST	81.60	
	120983	AA398209	Hs.97587	EST		81.10
	121034	AL389951	Hs.271623	nucleoporin 50kD		66.20
	121423	AW973352	Hs.290585	ESTs	64.40	
75	122553	AA451884	Hs.190121	ESTs		60.40
	122946	AI718702	Hs.308026	major histocompatibility complex, class	188.60	
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H		80.20
	124472	N52517	Hs.102670	EST	71.00	
80	124526	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci		104.90
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A		72.00
	125731	R61771	Hs.26912	ESTs		69.90
	125747	NM_002884	Hs.865	RAP1A, member of RAS oncogene family	69.00	
	126020	H79863	Hs.114243	ESTs		62.40
	126547	U47732	Hs.84072	transmembrane 4 superfamily member 3		62.80
85	126966	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra		60.10



	127472	AA761378	Hs.192013	ESTs	70.20	
	127610	AA960867	Hs.150271	ESTs, Highly similar to unnamed protein	64.00	
	127742	AW293496	Hs.180138	ESTs	85.20	
	127987	AI022103	Hs.124511	ESTs	96.60	
5	128233	AW889132	Hs.11916	ribokinase		78.90
	128420	AA650274	Hs.41296	fibronectin leucine rich transmembrane p		106.90
	128766	AW160432	Hs.296460	craniofacial development protein 1	66.80	
	129014	AW935187	Hs.170162	KIAA1357 protein		58.53
	129215	AB040930	Hs.126085	KIAA1497 protein	64.20	
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 18)	63.80	
	130385	AW067800	Hs.155223	stanniocalcin 2		139.60
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)		64.60
	131025	AB040900	Hs.6189	KIAA1467 protein	64.40	
	131241	BE501914	Hs.24654	Homo sapiens cDNA FLJ11640 fis, clone HE	76.20	
15	131775	AB014548	Hs.31921	KIAA0648 protein	97.80	
	132240	AB018324	Hs.42676	KIAA0781 protein		71.00
	132856	NM_001448	Hs.58367	glypican 4		88.40
	132977	AA093322	Hs.301404	RNA binding motif protein 3	133.20	
	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		59.30
20	133818	AI110684	Hs.7645	fibrinogen, B beta polypeptide	341.00	
	134264	AF149297	Hs.8087	NAG-5 protein		64.30
	134265	M83772	Hs.80876	flavin containing monooxygenase 3		232.53
	134346	X84002	Hs.82037	TATA box binding protein (TBP)-associate	66.00	
	134395	AA456539	Hs.8262	lysosomal-associated membrane protein 2		75.80
25	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su		108.30
	135056	N75765	Hs.93765	lipoma HMGIC fusion partner	71.40	
	135309	AI564123	Hs.42500	ADP-ribosylation factor-like 5	70.40	

30 TABLE 7B shows the accession numbers for those primekeys lacking unigenelD's for Table 7A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

35 Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

	Pkey	CAT number	Accessions
40	103207	30635_-4	X72790
	106566	120358_1	BE298210 AI672315 AW086489 BE298417 AA455921 AA902537 BE327124 R14963 AA085210 AW274273 AI333584 AI369742 AI039658 AI885095 AI476470 AI287650 AI885299 AI985381 AW592624 AW340136 AI266556 AA456390 AI310815 AA484951
45	116571	genbank_D45652	D45652
	118466	genbank_N66741	N66741
	101046	entrez_K01160	K01160
	101941	entrez_S77583	S77583
	103351	entrez_X89211	X89211
50	123130	genbank_AA487200	AA487200



Table 8A shows 1720 genes either up or down-regulated in lung tumors or chronically diseased lung relative to a broad collection of over 40 distinct normal body tissues. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 39494 probesets on the Eos/Affymetrix Hu02 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

Pkey: Unique Eos probeset identifier number

ExAccn: Exemplar Accession number, Genbank accession number

UnigeneID: Unigene number

Unigene Title: Unigene gene title

R1: 70th percentile of AI for lung tumors divided by 90th percentile of AI for normal lung

R2: 70th percentile of AI for chronically diseased lung divided by 90th percentile of AI for normal lung

	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
5						
10						
15	300097	AI916973	Hs.213603	ESTs	5.46	4.69
	300117	AW189787	Hs.147474	ESTs	0.58	0.56
	300197	AI686661	Hs.218286	ESTs	4.26	5.44
	300201	AI308300		gb:ta90c06.x1 NCI_CGAP_Bm20 Homo sapien	0.62	0.83
20	300225	AI989963	Hs.197505	ESTs	1.68	1.75
	300247	AW274682	Hs.161394	ESTs	1.08	2.28
	300256	AI469095	Hs.298241	Transmembrane protease, serine 3	0.86	1.00
	300337	AI707881	Hs.202090	ESTs	5.80	9.09
	300362	Z42308		gb:HSC0FB121 normalized infant brain cDN	4.18	12.78
25	300374	AI859947	Hs.314158	ESTs	2.99	4.38
	300387	AW270150	Hs.254516	ESTs	1.50	2.53
	300440	AI421541	Hs.146164	ESTs	3.98	5.25
	300441	R10367	Hs.307921	EST, Weakly similar to Z232_HUMAN ZINC F	3.18	6.80
	300449	AI362967	Hs.132221	hypothetical protein FLJ12401	0.43	0.62
30	300469	AW135830	Hs.233955	hypothetical protein FLJ20401	0.16	0.83
	300552	X85711	Hs.21838	hypothetical protein FLJ11191	4.10	9.75
	300627	W27363		gb:ab37d01.r1 Stratagene HeLa cell s3 93	4.60	12.60
	300630	AW118822	Hs.128757	ESTs	2.91	5.86
	300716	AI216113	Hs.126280	hypothetical protein FLJ23393	1.00	0.92
35	300738	AI623332	Hs.130541	KIAA1542 protein	1.82	1.71
	300777	AA235361	Hs.96840	KIAA1527 protein	4.48	8.22
	300790	AI492471	Hs.188270	ESTs	1.29	1.18
	300832	AI688147	Hs.220615	ESTs, Weakly similar to T03829 transcrip	5.51	8.56
	300836	Z44942	Hs.22958	calcium channel alpha2-delta3 subunit	4.90	6.34
40	300838	AI582897	Hs.192570	hypothetical protein FLJ22028	1.70	2.81
	300878	AW449802	Hs.285901	Homo sapiens cDNA FLJ20428 fis, clone KA	4.56	7.91
	300897	AI890356	Hs.127804	ESTs, Weakly similar to T17233 hypotheti	2.23	1.58
	300926	AA504860		gb:ab03a10.s1 Stratagene fetal retina 93	2.13	3.50
	300960	AI041019	Hs.152454	ESTs	2.74	4.46
45	300961	AW204069	Hs.312716	ESTs, Weakly similar to unnamed protein	1.00	1.00
	300962	AA593373	Hs.293744	ESTs	1.46	1.51
	300967	AA565209	Hs.269439	ESTs	0.39	1.30
	300987	AW450840	Hs.148590	ESTs, Weakly similar to AF208846 1 BM-00	1.49	1.08
	300988	AI927208	Hs.208952	ESTs	0.16	0.37
50	301050	AW136973	Hs.288516	ESTs, Weakly similar to S69890 mitogen i	3.23	1.94
	301098	AA677570	Hs.185918	ESTs	6.76	14.28
	301157	AA729905	Hs.231916	ESTs	3.16	8.85
	301162	AI142118	Hs.129004	ESTs	1.68	7.18
	301170	AA737594	Hs.247606	ESTs	4.40	6.42
55	301192	AI808751	Hs.121188	ESTs	6.38	11.59
	301193	AA758115	Hs.128350	ESTs, Weakly similar to JC5423 2-hydroxy	4.35	7.78
	301267	AW297762	Hs.255690	ESTs	1.56	1.61
	301281	AA843986	Hs.190586	ESTs	2.19	1.78
	301341	AI819198	Hs.208229	ESTs	0.76	0.76
60	301382	AA912839	Hs.163369	ESTs	1.00	1.81
	301407	AW450466	Hs.126830	ESTs	1.48	1.51
	301452	AA975688	Hs.159955	ESTs	0.51	1.46
	301483	AW272467	Hs.254655	Untitled	2.40	5.02
	301494	AI678034	Hs.131099	ESTs	2.79	3.41
65	301521	AI733621	Hs.133011	zinc finger protein 117 (HPF9)	0.67	0.67
	301531	AI077462	Hs.134084	ESTs	2.52	3.76
	301580	AI878959	Hs.73737	splicing factor, arginine/serine-rich 1	7.41	11.92
	301676	Z43570	Hs.27453	ESTs, Moderately similar to G01251 Rar p	8.31	10.70
	301690	F05865	Hs.108323	ubiquitin-conjugating enzyme E2E 2 (homo	2.70	4.22
70	301718	F07744	Hs.7987	DKFZP434F162 protein	4.20	8.78
	301799	AA384252	Hs.286132	D15F37 (pseudogene)	5.93	7.04
	301804	AA581004	Hs.62180	anillin (Drosophila Scraps homolog), act	1.70	0.76
	301822	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	1.58	1.36
	301846	R20002	Hs.6823	hypothetical protein FLJ10430	1.00	1.00
75	301868	T71508	Hs.13861	ESTs, Weakly similar to pH sensitive max	2.88	5.49
	301882	T78054		gb:yc97g09.r1 Soares infant brain 1NIB H	2.28	3.80
	301905	AI991127	Hs.117202	ESTs	1.00	1.00
	301948	AA344647	Hs.116724	aldo-keto reductase family 1, member B11	5.28	2.28
	301960	AW070252	Hs.27973	KIAA0874 protein	5.38	6.48
80	302011	T91418	Hs.125156	transcriptional adaptor 2 (ADA2, yeast,	3.03	3.42
	302016	N40834	Hs.23495	hypothetical protein FLJ11252	1.00	1.25
	302041	NM_001501	Hs.129715	gonadotropin-releasing hormone 2	0.71	0.99
	302072	AJ238381	Hs.132576	paired box gene 9	1.60	1.71
	302094	AI286176	Hs.6786	ESTs	0.52	1.20
85	302095	AW044300	Hs.137506	Homo sapiens BAC clone RP11-120J2 from 7	2.75	4.93
	302148	AW269618	Hs.23244	ESTs	3.04	3.87



	302155	AI088485	Hs.144759	ESTs	0.45	1.15
	302201	AJ006276	Hs.159003	transient receptor potential channel 6	0.33	0.84
	302202	AF097159	Hs.159140	UDP-Gal:betaGlcNAc beta 1,4- galactosylt	0.52	0.94
5	302206	AI937193	Hs.41143	phosphoinositide-specific phospholipase	2.76	3.65
	302209	AF047445	Hs.159297	killer cell lectin-like receptor subfami	1.00	1.00
	302235	AL049987	Hs.166361	Homo sapiens mRNA; cDNA DKFZp564F112 (fr	1.68	1.50
	302290	AL117607	Hs.175563	Homo sapiens mRNA; cDNA DKFZp564N0763 (f	1.00	2.11
	302328	AA354849	Hs.23240	Homo sapiens cDNA FLJ13496 fis, clone PL	9.38	13.08
10	302346	AL039101	Hs.194625	dynein, cytoplasmic, light intermediate	3.27	7.24
	302360	AJ010901	Hs.198267	mucin 4, tracheobronchial	2.54	1.88
	302384	Y08982	Hs.202676	synaptonemal complex protein 2	1.00	0.91
	302406	U86751	Hs.211956	CD3-epsilon-associated protein; antisens	2.63	2.67
	302409	AF155156	Hs.218028	adaptor-related protein complex 4, epsil	5.82	9.34
15	302423	AB028977	Hs.225974	KIAA1054 protein	3.66	3.18
	302432	AL080068	Hs.272534	Homo sapiens mRNA; cDNA DKFZp564J062 (fr	2.44	6.77
	302435	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	0.44	0.84
	302437	AB024730	Hs.227473	UDP-N-acetylglucosamine:a-1,3-D-mannosid	4.18	5.64
	302455	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	1.85	0.92
20	302472	AA317451	Hs.6335	SWI/SNF related, matrix associated, acti	2.04	2.13
	302476	AF182294	Hs.241578	U6 snRNA-associated Sm-like protein LSm8	1.44	1.89
	302489	T80660	Hs.230424	Homo sapiens cDNA FLJ13540 fis, clone PL	0.51	1.10
	302490	AA885502	Hs.187032	ESTs	2.64	4.87
	302562	AJ005585	Hs.48956	gap junction protein, beta 6 (connexin 3	5.34	2.68
25	302566	AA085996	Hs.248572	hypothetical protein FLJ22965	1.00	1.21
	302630	AB029488	Hs.272100	SMS3 protein	0.52	1.24
	302634	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	1.00	1.00
	302638	AA463798	Hs.102696	MCT-1 protein	1.58	1.02
	302647	X57723	Hs.198273	NADH dehydrogenase (ubiquinone) 1 beta s	2.72	6.85
30	302655	AJ227892	Hs.146274	ESTs	1.00	4.32
	302656	AW293005	Hs.70704	Homo sapiens, clone IMAGE:2823731, mRNA,	2.97	0.93
	302668	AA580691	Hs.180789	S164 protein	0.80	0.95
	302679	H65022		gb:yu66g11.r1 Weizmann Olfactory Epithel	1.68	5.04
	302680	AW192334	Hs.38218	ESTs	2.70	7.98
35	302697	AJ001408		gb:Homo sapiens mRNA for immunoglobulin	4.25	8.13
	302705	U09060		gb:Human immunoglobulin heavy chain, V-r	3.91	8.68
	302711	L08442		gb:Human autonomously replicating sequen	2.20	2.73
	302719	W69724	Hs.288959	hypothetical protein FLJ20920	0.54	1.02
	302742	L12069		gb:Homo sapiens (clone WR4.10VH) anti-th	4.28	11.57
40	302755	AW384815	Hs.149208	KIAA1555 protein	1.57	2.38
	302771	H98476	Hs.42522	ESTs	2.94	4.68
	302789	AJ245067		gb:Homo sapiens mRNA for immunoglobulin	3.49	6.31
	302795	AJ245313	Hs.272838	hypothetical protein FLJ10494	0.80	2.74
	302802	Y08250		gb:H.sapiens mRNA for variable region of	1.13	0.77
45	302803	AA442824	Hs.293961	ESTs, Moderately similar to putative DNA	3.14	10.68
	302812	N31301	Hs.152664	hypothetical protein FLJ20051	3.04	8.24
	302847	X98940		gb:H.sapiens rearranged Ig heavy chain (	1.80	1.92
	302885	AL137763	Hs.132127	hypothetical protein LOC57822	1.00	1.00
	302943	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	0.53	0.67
50	302977	AW263124	Hs.315111	hypothetical protein FLJ12894	2.45	2.62
	303006	AF078950	Hs.24139	Homo sapiens cDNA: FLJ23137 fis, clone L	4.88	8.61
	303011	AF090405		gb:Homo sapiens clone 2A1 scFV antibody	1.41	1.86
	303013	F07898	Hs.288968	RAB22A, member RAS oncogene family	1.51	1.19
	303061	AF151882	Hs.27693	peptidylprolyl isomerase (cyclophilin)-I	0.72	0.76
55	303077	AF163305		gb:H.sapiens T-cell receptor mRNA	1.17	3.90
	303090	AA443259	Hs.146286	kinesin family member 13A	4.08	6.46
	303091	AF192913	Hs.130683	zinc finger protein 180 (HHZ168)	2.50	4.37
	303094	AF195513	Hs.278953	Pur-gamma	5.38	8.38
	303095	AF202051	Hs.134079	NM23-H8	3.26	4.08
60	303131	AW081061	Hs.103180	DC2 protein	2.02	1.83
	303195	AA082211	Hs.233936	myosin, light polypeptide, regulatory, n	1.32	3.95
	303196	AA082298	Hs.59710	ESTs	0.77	0.53
	303216	AA581439	Hs.152328	ESTs	0.24	0.63
	303222	AA333538	Hs.204501	hypothetical protein FLJ10534	3.56	6.22
65	303234	AA132255	Hs.143951	ESTs	2.28	3.17
	303251	AW340037	Hs.115897	protocadherin 12	0.38	1.02
	303295	AA205625	Hs.208067	ESTs	2.30	1.00
	303297	T80072	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.86	4.48
	303316	AF033122	Hs.14125	p53 regulated PA26 nuclear protein	0.10	0.80
70	303467	AA398801	Hs.323397	ESTs	4.54	9.65
	303506	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	0.09	0.04
	303552	AA359799	Hs.224662	ESTs, Weakly similar to unnamed protein	1.00	1.72
	303598	AA382814		gb:EST96097 Testis I Homo sapiens cDNA 5	4.96	9.14
	303637	AF056083	Hs.24879	phosphatidic acid phosphatase type 2C	2.06	2.02
75	303655	AA504702	Hs.258802	ATPase, (Na+)/K+ transporting, beta 4 po	1.00	1.24
	303756	AI738488	Hs.115838	ESTs	1.08	1.43
	303856	AA968589	Hs.180532	glucose phosphate isomerase	1.76	1.31
	303893	N88597	Hs.113503	karyopherin (importin) beta 3	2.30	2.57
	303907	AW467774	Hs.171880	polymerase (RNA) II (DNA directed) polyp	3.10	5.79
80	303946	AW474196	Hs.306637	Homo sapiens cDNA FLJ12363 fis, clone MA	5.06	11.86
	303978	AW513315		gb:xo43c12.x1 NCI_CGAP_Ut1 Homo sapiens	5.14	7.31
	303981	AW513804	Hs.278834	ESTs, Weakly similar to ALU1_HUMAN ALU S	2.83	4.06
	303990	AW515465		gb:xu71a11.x1 NCI_CGAP_Kid8 Homo sapiens	1.15	2.35
	303998	AW516449		gb:xt68f05.x1 NCI_CGAP_Ut2 Homo sapiens	2.20	9.35
85	303999	AW516611		gb:xp70b11.x1 NCI_CGAP_Ov39 Homo sapiens	4.85	6.28
	304006	AW517947		gb:xt66h02.x1 NCI_CGAP_Ut2 Homo sapiens	3.21	4.07



	304008	AW518198	Hs.3297	ribosomal protein S27a	6.50	11.08
	304009	AW518206	Hs.181165	eukaryotic translation elongation factor	1.88	3.27
	304024	T03036		gb:FB21B7 Fetal brain, Stratagene Homo s	2.15	3.55
5	304026	T03160		gb:FB26F2 Fetal brain, Stratagene Homo s	5.88	11.80
	304028	T03266		gb:FB7C1 Fetal brain, Stratagene Homo sa	5.59	13.46
	304036	T16855	Hs.244621	ribosomal protein S14	6.55	14.43
	304046	T54803		gb:yb42d06.s1 Stratagene fetal spleen (9	6.18	12.19
	304061	T61521		gb:yb73g01.s1 Stratagene ovary (937217)	2.64	8.23
10	304063	T62536		gb:yc04c12.s1 Stratagene lung (937210) H	0.53	1.61
	304097	R25376	Hs.177592	ribosomal protein, large, P1	6.49	11.67
	304114	R78946		gb:yi87g02.s1 Soares placenta Nb2HP Homo	2.90	4.18
	304122	H28966		gb:ym31a06.s1 Soares infant brain 1NIB H	1.00	2.76
	304155	H68696		gb:yr78b06.s1 Soares fetal liver spleen	0.79	1.18
15	304203	N56929		gb:yy82d08.s1 Soares_multiple_sclerosis_	4.28	11.34
	304234	W81608		gb:zd88h06.s1 Soares_fetal_heart_NbHH19W	6.47	11.03
	304267	AA064862	Hs.73742	ribosomal protein, large, P0	1.34	1.16
	304270	AA069711	Hs.297753	vimentin	3.40	5.40
	304287	AA079286	Hs.78466	proteasome (prosome, macropain) 26S sub	2.93	4.42
20	304348	AA179868		gb:zp38g12.s1 Stratagene muscle 937209 H	3.98	10.96
	304415	AA290747	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	3.32	5.99
	304430	AA347682		gb:EST54044 Fetal heart II Homo sapiens	1.00	1.00
	304456	AA411240		gb:zv26g05.s1 Soares_NhHMPu_S1 Homo sapi	1.42	3.33
	304521	AA464716		gb:zx82c11.s1 Soares ovary tumor NbHOT H	2.18	1.15
25	304526	AA476427		gb:zx02c05.s1 Soares_total_fetus_Nb2HF8_	5.38	14.11
	304542	AA482602	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.16	8.23
	304546	AA486074	Hs.297681	serine (or cysteine) proteinase inhibito	0.55	1.20
	304607	AA513322		gb:nh85e08.s1 NCI_CGAP_Br1.1 Homo sapien	1.95	2.10
	304640	AA524440	Hs.111334	ferritin, light polypeptide	2.10	2.83
30	304650	AA527489	Hs.3463	ribosomal protein S23	3.33	12.62
	304735	AA576453		gb:nm75h11.s1 NCI_CGAP_Co9 Homo sapiens	1.33	0.88
	304760	AA580401		gb:nn13g09.s1 NCI_CGAP_Co12 Homo sapiens	3.68	8.14
	304849	AA588157	Hs.13801	KIAA1685 protein	2.77	3.70
	304917	AA602685	Hs.284136	PRO2047 protein	7.16	11.01
35	304921	AA603092	Hs.297753	vimentin	2.47	4.24
	304966	AA613893	Hs.282435	ESTs	6.78	11.66
	304987	AA618044	Hs.300697	immunoglobulin heavy constant gamma 3 (G	0.90	1.23
	305016	AA626876		gb:zu89h06.s1 Soares_testis_NHT Homo sap	6.46	10.17
	305034	AA630128		gb:ab99c04.s1 Stratagene lung (937210) H	1.00	1.00
40	305072	AA641012		gb:nr72a12.s1 NCI_CGAP_Pr24 Homo sapiens	5.68	11.59
	305111	AA644187	Hs.303405	ESTs	1.48	1.37
	305148	AA654070		gb:nt01g08.s1 NCI_CGAP_Lym3 Homo sapiens	1.76	4.61
	305159	AA659166	Hs.275668	EST, Weakly similar to EF1D_HUMAN ELONG	1.00	2.15
	305190	AA665955		gb:ag57d12.s1 Gessler Wilms tumor Homo s	5.31	8.14
45	305232	AA670052	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	0.78	1.18
	305235	AA670480		gb:ag37e01.s1 Jia bone marrow stroma Hom	3.11	8.66
	305245	AA676695	Hs.81328	nuclear factor of kappa light polypeptid	4.38	7.53
	305312	AA700201		gb:zj44f07.s1 Soares_fetal_liver_spleen_	2.13	2.66
	305322	AA701597	Hs.163019	EST	1.20	1.40
50	305394	AA720942	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.16	0.68
	305413	AA724659		gb:ai10f08.s1 Soares_parathyroid_tumor_N	5.86	9.87
	305447	AA737856		gb:nx10c08.s1 NCI_CGAP_GC3 Homo sapiens	2.21	2.86
	305476	AA745664	Hs.287445	hypothetical protein FLJ11726	3.36	6.54
	305483	AA748030	Hs.303512	EST	1.00	2.02
55	305528	AA769156		gb:nz12e05.s1 NCI_CGAP_GCB1 Homo sapiens	6.44	9.10
	305612	AA782347	Hs.272572	hemoglobin, alpha 2	0.19	0.79
	305614	AA782866		gb:aj09h02.s1 Soares_parathyroid_tumor_N	1.00	1.00
	305616	AA782884	Hs.275865	ribosomal protein S18	7.57	10.20
	305637	AA806124		gb:oe29a12.s1 NCI_CGAP_Pr25 Homo sapiens	4.78	12.42
60	305639	AA806138		gb:oe29c12.s1 NCI_CGAP_Pr25 Homo sapiens	0.89	0.70
	305650	AA807709		gb:nw31e04.s1 NCI_CGAP_GCB0 Homo sapiens4.49	4.49	8.71
	305690	AA813477		gb:ai67a05.s1 Soares_testis_NHT Homo sap	4.91	9.40
	305726	AA828156	Hs.73742	ribosomal protein, large, P0	0.19	0.81
	305728	AA828209		gb:of34a02.s1 NCI_CGAP_Kid6 Homo sapiens	5.12	9.29
65	305759	AA835353		gb:ak72b06.s1 Barstead spleen HPLRB2 Hom	1.66	4.11
	305792	AA845256		gb:ak84a08.s1 Barstead spleen HPLRB2 Hom	2.34	4.25
	305864	AA864374	Hs.73742	ribosomal protein, large, P0	0.30	1.40
	305901	AA872968		gb:oh63h08.s1 NCI_CGAP_Kid5 Homo sapiens	2.10	5.21
	305910	AA875981		gb:nx21h02.s1 NCI_CGAP_GC3 Homo sapiens	0.32	1.01
70	306015	AA897116		gb:am08b07.s1 Soares_NFL_T_GBC_S1 Homo s1.56	1.56	1.12
	306017	AA897221	Hs.109058	ribosomal protein S6 kinase, 90kD, polyp	5.21	7.90
	306020	AA897630	Hs.130027	EST	1.96	6.59
	306063	AA906316		gb:ok03g03.s1 Soares_NFL_T_GBC_S1 Homo s	7.38	20.69
	306065	AA906725		gb:ok78g02.s1 NCI_CGAP_GC4 Homo sapiens	7.19	13.48
75	306104	AA910956		gb:ok85h11.s1 NCI_CGAP_Kid3 Homo sapiens	6.50	9.13
	306109	AA911861		gb:og21a07.s1 NCI_CGAP_PNS1 Homo sapiens	4.21	5.25
	306148	AA917409	Hs.288036	tRNA isopentenylpyrophosphate transferas	2.20	2.70
	306242	AA932805		gb:oo60g04.s1 NCI_CGAP_Lu5 Homo sapiens	2.84	5.35
	306288	AA936900		gb:oi53h05.s1 NCI_CGAP_HN3 Homo sapiens	1.60	1.12
80	306325	AA953072	Hs.210546	interleukin 21 receptor	1.65	2.26
	306353	AA961382	Hs.275865	ribosomal protein S18	3.78	6.32
	306375	AA968650	Hs.276018	EST, Moderately similar to JC4662 ribos	4.30	5.74
	306396	AA970223		gb:op09d05.s1 NCI_CGAP_Kid6 Homo sapiens	0.95	2.45
	306428	AA975110	Hs.191228	hypothetical protein FLJ20284	3.19	4.10
	306442	AA976899		gb:oo35e09.s1 NCI_CGAP_GC4 Homo sapiens	4.67	7.44
85	306446	AA977348		gb:oo72e12.s1 NCI_CGAP_Kid6 Homo sapiens	3.92	6.27



5	306458	AA978186		gb:op33c06.s1 Soares_NFL_T_GBC_S1 Homo s	3.35	5.77
	306467	AA983508	Hs.163593	ribosomal protein L18a	3.72	5.37
	306510	AA988546		gb:or84d07.s1 NCI_CGAP_Lu5 Homo sapiens	1.00	1.00
	306555	AA994304	Hs.276083	EST, Weakly similar to RL23_HUMAN 60S R	6.61	10.91
	306557	AA994530		gb:ou57e08.s1 NCI_CGAP_Br2 Homo sapiens	16.20	31.83
10	306572	AA995686		gb:os25c12.s1 NCI_CGAP_Kid5 Homo sapiens	2.51	6.52
	306582	AA996248		gb:os18c10.s1 NCI_CGAP_Kid5 Homo sapiens	1.42	3.13
	306598	AI000320	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.91	8.68
	306605	AI000497	Hs.119500	ribosomal protein, large P2	1.96	8.60
	306656	AI004024		gb:ou11b07.x1 Soares_NFL_T_GBC_S1 Homo s	0.11	0.45
15	306676	AI005603	Hs.284136	PRO2047 protein	9.56	17.28
	306686	AI015615		gb:ov29f10.x1 Soares_testis_NHT Homo sap	1.86	3.60
	306702	AI022565	Hs.307670	EST	1.47	1.19
	306728	AI027359	Hs.272572	hemoglobin, alpha 2	1.28	2.83
	306751	AI032589		gb:ow70h12.s1 Soares_fetal_liver_spleen_	3.91	5.21
20	306767	AI038963	Hs.249118	ESTs	3.33	6.06
	306892	AI092465		gb:qa75h12.x1 Soares_fetal_heart_NbHH19W	3.77	7.46
	306897	AI093967		gb:qa33c06.s1 Soares_NhHMPu_S1 Homo sapi	2.12	2.85
	306956	AI125111		gb:am66f03.s1 Barstead spleen HPLRB2 Hom	6.10	10.52
	306958	AI125152		gb:am55e09.x1 Johnston frontal cortex Ho	1.72	1.56
25	307035	AI142774	Hs.119122	ribosomal protein L13a	2.00	4.70
	307041	AI144243		gb:qb85b12.x1 Soares_fetal_heart_NbHH19W	9.12	12.56
	307091	AI167439		gb:ox70h06.s1 Soares_NhHMPu_S1 Homo sapi	4.88	8.52
	307181	AI189251		gb:qc99g06.x1 Soares_pregnant_uterus_NbH	3.55	6.44
	307297	AI205798	Hs.111334	ferritin, light polypeptide	2.46	4.65
30	307317	AI208303	Hs.147333	EST	5.64	10.13
	307327	AI214142	Hs.246381	CD68 antigen	3.18	5.15
	307382	AI223158	Hs.147885	ESTs	2.02	3.73
	307410	AI241715	Hs.77039	ribosomal protein S3A	0.72	0.48
	307415	AI242118		gb:qh92b02.x1 Soares_NFL_T_GBC_S1 Homo s	2.38	3.51
35	307423	AI243206	Hs.179573	collagen, type I, alpha 2	2.60	5.44
	307426	AI243364		gb:qh30g11.x1 Soares_NFL_T_GBC_S1 Homo s	3.18	7.67
	307517	AI275055		gb:ql72d03.x1 Soares_NhHMPu_S1 Homo sapi	1.00	1.00
	307551	AI281556		gb:qu52f11.x1 NCI_CGAP_Lym6 Homo sapiens	3.40	11.20
	307561	AI282207		gb:qp65a12.x1 Soares_fetal_lung_NbHL19W	4.74	15.51
40	307608	AI290295		gb:qm01f02.x1 Soares_NhHMPu_S1 Homo sapi	3.50	7.19
	307657	AI306428	Hs.298262	ribosomal protein S19	1.76	2.44
	307691	AI318285		gb:tb17b01.x1 NCI_CGAP_Ov37 Homo sapiens	1.59	1.31
	307701	AI318583	Hs.276672	EST, Weakly similar to RL6_HUMAN 60S RI	1.90	2.13
	307718	AI333406	Hs.83753	small nuclear ribonucleoprotein polypept	0.45	0.99
45	307730	AI336092		gb:qt43b07.x1 Soares_fetal_lung_NbHL19W	1.51	0.99
	307760	AI342387		gb:qt27f07.x1 Soares_pregnant_uterus_NbH	1.00	1.00
	307764	AI342731		gb:qo26a07.x1 NCI_CGAP_Lu5 Homo sapiens	4.52	12.58
	307783	AI347274		gb:tc05d02.x1 NCI_CGAP_Co16 Homo sapiens	1.42	1.00
	307796	AI350556		gb:qt18f09.x1 NCI_CGAP_GC4 Homo sapiens	6.57	9.61
50	307807	AI351799		gb:qt09d02.x1 NCI_CGAP_GC4 Homo sapiens	3.38	7.68
	307808	AI351826		gb:qt09g03.x1 NCI_CGAP_GC4 Homo sapiens	0.33	0.86
	307820	AI355761		gb:qt94a11.x1 NCI_CGAP_Co14 Homo sapiens	7.94	21.57
	307830	AI358722	Hs.276737	EST, Weakly similar to R5HU22 ribosomal	2.05	3.32
	307852	AI365541		gb:qz08g05.x1 NCI_CGAP_CLL1 Homo sapiens	3.18	5.21
55	307902	AI380462		gb:tg02h05.x1 NCI_CGAP_CLL1 Homo sapiens	3.13	4.99
	307997	AI434512	Hs.181165	eukaryotic translation elongation factor	1.00	3.01
	308002	AI435240	Hs.283442	ESTs	5.86	12.64
	308011	AI439473		gb:ti60a08.x1 NCI_CGAP_Lym12 Homo sapien	3.79	5.83
	308023	AI452732	Hs.251577	hemoglobin, alpha 1	0.38	0.88
60	308041	AI458824	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.36	6.06
	308059	AI468938	Hs.276877	EST, Weakly similar to RL10_HUMAN 60S R	1.80	1.98
	308085	AI474135	Hs.181165	eukaryotic translation elongation factor	3.38	4.14
	308101	AI475950	Hs.181165	eukaryotic translation elongation factor	1.30	3.87
	308106	AI476803		gb:lj77e12.x1 Soares_NSF_F8_9W_OT_PA_P_S2.38		8.72
65	308122	AI480123	Hs.309411	EST	2.70	3.86
	308154	AI500600		gb:tn93d08.x1 NCI_CGAP_Ui2 Homo sapiens	0.66	1.33
	308171	AI523632	Hs.298766	ESTs, Weakly similar to schlafen4 [M.mu	2.48	4.86
	308211	AI557029	Hs.278572	anaplastic lymphoma kinase (Ki-1)	2.43	2.14
	308213	AI557041		gb:PT2.1_12_E04.r tumor2 Homo sapiens cD	3.34	3.79
70	308216	AI557135		gb:PT2.1_13_H06.r tumor2 Homo sapiens cD	4.61	4.78
	308219	AI557246		gb:PT2.1_15_D07.r tumor2 Homo sapiens cD	4.87	7.94
	308271	AI567844	Hs.252259	ribosomal protein S3	2.40	6.35
	308319	AI583983	Hs.181165	eukaryotic translation elongation factor	2.45	3.33
	308362	AI613519	Hs.105749	KIAA0553 protein	1.24	1.41
75	308413	AI636253	Hs.196511	ESTs	3.16	4.82
	308450	AI660860	Hs.96840	KIAA1527 protein	1.79	2.68
	308464	AI672425	Hs.277117	EST, Moderately similar to I38055 myosi	4.87	8.27
	308588	AI718299		gb:as51g12.x1 Barstead aorta HPLRB6 Homo	3.90	5.64
	308599	AI719893		gb:as47d07.x1 Barstead aorta HPLRB6 Homo	3.32	5.12
80	308615	AI738593	Hs.101774	hypothetical protein FLJ23045	3.11	2.36
	308643	AI745040		gb:tr19a12.x1 NCI_CGAP_Ov23 Homo sapiens	3.98	3.69
	308673	AI760864		gb:wi09c10.x1 NCI_CGAP_CLL1 Homo sapiens	0.82	0.99
	308697	AI767143		gb:wi97a07.x1 NCI_CGAP_Kid12 Homo sapien	2.76	5.59
	308762	AI807405	Hs.259408	ESTs	3.17	6.30
85	308778	AI811109		gb:tr04c11.x1 NCI_CGAP_Ov23 Homo sapiens	1.00	1.00
	308782	AI811767	Hs.2186	eukaryotic translation elongation factor	2.94	5.15
	308808	AI818289		gb:wk52c01.x1 NCI_CGAP_Pr22 Homo sapiens	4.41	8.34
	308823	AI824118	Hs.217493	annexin A2	1.85	1.92
	308875	AI832332		gb:at48g03.x1 Barstead colon HPLRB7 Homo	2.52	3.80



5	308879	AI832763	Hs.75968	thymosin, beta 4, X chromosome	3.38	7.96
	308886	AI833240		gb:at76d10.x1 Barstead colon HPLRB7 Homo	3.06	2.65
	308898	AI858845		gb:wl32d10.x1 NCI_CGAP_Ut1 Homo sapiens	2.45	3.44
	308934	AI865023	Hs.177	phosphatidylinositol glycan, class H	4.14	6.76
	308966	AI870704		gb:wl47h01.x1 NCI_CGAP_Ut1 Homo sapiens	1.00	1.00
10	308979	AI873111		gb:wl52h05.x1 NCI_CGAP_Brn25 Homo sapien	7.15	11.10
	309045	AI910902		gb:tg39f01.x1 NCI_CGAP_Ut1 Homo sapiens	0.61	0.59
	309051	AI911975		gb:wd78d01.x1 NCI_CGAP_Lu24 Homo sapiens	1.78	4.42
	309069	AI917366	Hs.78202	SWI/SNF related, matrix associated, act	3.27	5.88
	309083	AI922426	Hs.119598	ribosomal protein L3	2.39	3.34
15	309105	AI925503	Hs.265884	ESTs	5.54	17.78
	309122	AI928178		gb:wo95a11.x1 NCI_CGAP_Kid11 Homo sapien	1.00	2.92
	309128	AI928816	Hs.180842	ribosomal protein L13	1.38	5.55
	309164	AI937761		gb:wp84b09.x1 NCI_CGAP_Brn25 Homo sapien	2.43	3.11
	309177	AI951118		gb:wx63g05.x1 NCI_CGAP_Br18 Homo sapiens	0.81	0.97
20	309288	AI991525	Hs.299426	ESTs	4.86	7.46
	309299	AW003478		gb:wq66c06.x1 NCI_CGAP_GC6 Homo sapiens	4.36	9.43
	309303	AW004823		gb:ws93a08.x1 NCI_CGAP_Co3 Homo sapiens	2.88	7.54
	309411	AW085201	Hs.244144	EST	4.30	7.14
	309437	AW090702	Hs.278242	tubulin, alpha, ubiquitous	2.49	3.11
25	309459	AW117645	Hs.65114	keratin 18	2.88	4.55
	309476	AW129368		gb:xe14b05.x1 NCI_CGAP_Ut4 Homo sapiens	2.08	6.60
	309499	AW136325	Hs.279771	Homo sapiens clone PP1596 unknown mRNA	2.82	3.55
	309529	AW150807	Hs.181357	laminin receptor 1 (67kD, ribosomal pro	4.78	3.95
	309532	AW151119		gb:xg33e10.x1 NCI_CGAP_Ut1 Homo sapiens	1.18	4.40
30	309626	AW192004	Hs.297681	serine (or cysteine) proteinase inhibit	4.46	12.06
	309641	AW194230	Hs.253100	EST, Moderately similar to GHU Ig gamm	1.47	1.39
	309675	AW205681	Hs.253506	EST, Moderately similar to ATPN_HUMAN A	5.68	15.20
	309693	AW237221	Hs.181357	laminin receptor 1 (67kD, ribosomal prot	1.00	1.00
	309695	AW238011	Hs.295605	mannosidase, alpha, class 2A, member 2	5.45	9.61
35	309700	AW241170	Hs.179661	tubulin, beta polypeptide	1.41	1.25
	309747	AW264889		gb:xq36h02.x1 NCI_CGAP_Lu28 Homo sapiens	5.00	8.35
	309769	AW272346		gb:xs13c10.x1 NCI_CGAP_Kid11 Homo sapien	5.76	11.90
	309782	AW275156	Hs.156110	immunoglobulin kappa constant	0.42	0.69
	309783	AW275401	Hs.254798	EST	1.00	4.11
40	309799	AW276964		gb:xp58h01.x1 NCI_CGAP_Ov39 Homo sapiens	1.68	1.44
	309866	AW299916		gb:xs44c01.x1 NCI_CGAP_Kid11 Homo sapien	3.02	5.04
	309903	AW339071	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.05	1.18
	309923	AW340684		gb:hd05g08.x1 Soares_NFL_T_GBC_S1 Homo s	2.30	3.67
	309928	AW341418		gb:hd08c03.x1 Soares_NFL_T_GBC_S1 Homo s	7.41	13.71
45	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s	1.20	12.70
	309933	AW341936		gb:hb73f10.x1 NCI_CGAP_Ut2 Homo sapiens	4.90	18.29
	309964	AW449111	Hs.257111	hypothetical protein MGC3265	1.99	3.07
	310002	AI439096	Hs.323079	Homo sapiens mRNA; cDNA DKFZp564P116 (fr	0.20	0.47
	310096	AW136822	Hs.172824	ESTs, Weakly similar to B48013 proline-r	1.51	1.22
50	310098	AI685841	Hs.161354	ESTs	0.31	0.76
	310109	AI203094	Hs.148633	ESTs	2.06	5.83
	310112	AW197233	Hs.147253	ESTs	2.92	3.55
	310115	AI611317	Hs.223796	ESTs	1.25	0.84
	310121	AW195642	Hs.148901	ESTs	1.00	2.71
55	310146	AI206614	Hs.197422	ESTs	9.50	15.31
	310193	AI627653	Hs.147562	ESTs	2.85	4.18
	310255	AW450439	Hs.153378	ESTs	4.26	10.63
	310261	AI240483	Hs.201217	ESTs	3.28	4.40
	310264	AI915771	Hs.74170	metallothionein 1E (functional)	0.26	0.86
60	310275	AI242102	Hs.213636	ESTs	5.43	8.19
	310282	AI243332	Hs.156055	ESTs	3.15	8.06
	310290	AW013815	Hs.149103	ESTs	2.19	3.12
	310333	AI253200	Hs.145402	ESTs	1.17	1.91
	310346	AI261340	Hs.145517	ESTs	4.81	9.95
65	310385	AI263392	Hs.156151	ESTs	5.96	7.79
	310443	AW119018	Hs.164231	ESTs	2.90	4.63
	310444	AW196632	Hs.252956	ESTs	0.85	1.01
	310446	AI275715	Hs.145926	ESTs	2.18	3.85
	310468	AI984074	Hs.196398	ESTs	3.39	5.19
70	310477	AI948801	Hs.171073	ESTs	1.00	1.00
	310512	AW275603	Hs.200712	ESTs	3.87	8.12
	310514	AI681145	Hs.160724	ESTs	3.30	7.33
	310524	AW082270	Hs.12496	ESTs, Highly similar to AC004836 1 simil	0.72	1.44
	310547	AI302654	Hs.208024	ESTs	3.26	3.46
75	310584	AI653007	Hs.156304	ESTs	2.39	4.08
	310608	AI962234	Hs.196102	ESTs	5.60	6.49
	310624	AI341594		gb:Human endogenous retrovirus H proteas	4.91	9.09
	310636	AI814373	Hs.164175	ESTs	1.85	1.71
	310648	AI347863	Hs.156672	ESTs	0.17	0.69
80	310694	AI654370	Hs.157752	Homo sapiens mRNA full length insert cDN	5.40	13.22
	310695	AI472124	Hs.157757	ESTs	4.82	6.27
	310714	AI418446	Hs.157882	ESTs	1.76	3.51
	310722	AI989803	Hs.157289	ESTs	1.14	6.85
	310756	AI916560	Hs.158707	ESTs	8.46	13.01
85	310764	AI376769	Hs.167172	ESTs	4.76	7.37
	310848	AI459554	Hs.161286	ESTs	2.84	1.96
	310851	AW291714	Hs.221703	ESTs	1.00	2.32
	310854	AI421677	Hs.161332	ESTs	6.37	7.94
	310858	AI871000	Hs.161330	ESTs	6.07	9.84



	310864	AI924558	Hs.161399	ESTs	0.87	0.78
	310875	T47764	Hs.132917	ESTs	1.00	3.63
	310896	AW157731	Hs.270982	ESTs, Moderately similar to ALU7_HUMAN A	7.07	16.68
5	310922	AW195634	Hs.170401	ESTs	1.00	1.00
	310955	AI560210	Hs.263912	ESTs	10.08	17.66
	310957	AW190974	Hs.196918	ESTs	2.18	3.18
	311000	AI521830	Hs.171050	ESTs	3.06	6.64
	311012	AW298070	Hs.241097	ESTs	1.23	3.77
10	311034	AI564023	Hs.311389	ESTs, Moderately similar to PT0375 natur	2.44	2.09
	311074	AW290922	Hs.199848	ESTs	6.04	14.19
	311134	AI990849	Hs.196971	ESTs	3.54	6.96
	311174	AW450552	Hs.205457	periaxin	0.65	0.95
	311187	AI638374	Hs.224189	ESTs	2.46	2.78
15	311220	AI656040	Hs.196532	ESTs	1.10	2.52
	311230	AI989808	Hs.197663	ESTs	1.41	1.75
	311236	AI653378	Hs.197674	ESTs	2.18	2.11
	311242	AW016812	Hs.200266	ESTs	0.63	5.11
	311258	AI671221	Hs.199887	ESTs	1.00	1.41
20	311277	AW072813	Hs.270868	ESTs, Moderately similar to ALU4_HUMAN A	2.56	1.94
	311294	AA826425	Hs.291829	ESTs	1.04	2.69
	311308	F12664	Hs.49000	ESTs	1.96	6.70
	311351	AI682303	Hs.201274	ESTs	4.77	9.38
	311390	AW392997	Hs.202280	ESTs	2.80	6.06
25	311405	AW290961	Hs.201815	ESTs	3.80	11.66
	311409	AI698839		gb:wd31f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.84	6.94
	311420	AI936291	Hs.209867	ESTs	5.30	12.56
	311443	AI791521	Hs.192206	ESTs	4.39	6.09
	311467	AI934909	Hs.175377	ESTs	1.00	1.04
30	311479	AI933672	Hs.211399	ESTs	2.76	5.61
	311488	R57390	Hs.301064	arfaptin 1	2.50	5.73
	311495	AW300077	Hs.221358	ESTs	3.63	6.09
	311511	AW444568	Hs.210303	ESTs	2.00	2.87
	311534	AW130351	Hs.243549	ESTs	0.31	1.33
35	311537	AI805121	Hs.211828	ESTs	3.69	5.85
	311543	AI681360	Hs.201259	ESTs	1.73	1.34
	311551	AW449774	Hs.296380	POM (POM121 rat homolog) and ZP3 fusion	3.31	6.12
	311557	AI819230	Hs.211238	interleukin-1 homolog 1	1.00	1.00
	311558	Z44432	Hs.63128	KIAA1292 protein	2.25	3.41
40	311559	AW008271	Hs.265848	similar to rat myomegalin	2.68	5.90
	311563	AI922143	Hs.211334	ESTs	2.39	3.32
	311586	AI827834	Hs.211227	ESTs	2.47	3.85
	311616	AW450675	Hs.212709	ESTs	1.00	1.00
	311621	AI924307	Hs.213464	ESTs	4.16	6.74
45	311635	AI928456	Hs.213081	ESTs	2.17	3.76
	311668	AW193674	Hs.240044	ESTs	2.60	3.12
	311672	R11807	Hs.20914	hypothetical protein FLJ23056	2.79	5.18
	311683	AW183738	Hs.232644	ESTs	0.19	0.96
	311700	R49601	Hs.171495	retinoic acid receptor, beta	6.28	8.83
50	311714	AW131785	Hs.246831	ESTs, Weakly similar to CIKG_HUMAN VOLTA	5.00	8.17
	311735	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	0.96	0.72
	311743	T99079	Hs.191194	ESTs	1.00	1.95
	311783	AI682478	Hs.13528	hypothetical protein FLJ14054	0.16	0.77
	311785	AI056769	Hs.133512	ESTs	1.34	3.97
55	311799	AA780791	Hs.14014	ESTs, Weakly similar to KIAA0973 protein	8.52	13.32
	311819	AW265275	Hs.254325	ESTs	3.58	3.91
	311823	AI089422	Hs.131297	ESTs	1.40	1.72
	311877	AA349893	Hs.85339	G protein-coupled receptor 39	0.95	0.91
	311886	AA522738	Hs.132554	ESTs	0.88	0.87
60	311896	AW206447		gb:UI-H-BI1-afg-g-02-0-UI.s1 NCI_CGAP_Su	1.66	1.13
	311910	N28365	Hs.22579	Homo sapiens clone CDABP0036 mRNA sequen	1.66	2.30
	311923	T60843	Hs.189679	ESTs	0.42	2.63
	311933	AI597963	Hs.118726	ESTs	1.88	3.02
	311959	T67262	Hs.124733	ESTs	2.02	2.33
65	311960	AW440133	Hs.189690	ESTs	3.87	6.62
	311967	AI382726	Hs.182434	ESTs	5.80	8.14
	311975	AA804374	Hs.272203	Homo sapiens cDNA FLJ20843 fis, clone AD	0.98	3.26
	312005	T78450	Hs.13941	ESTs	0.12	1.39
	312028	T78886	Hs.284450	ESTs	3.78	4.92
70	312046	AI580018	Hs.268591	ESTs	4.11	7.32
	312056	T83748	Hs.268594	ESTs	2.36	3.08
	312064	AA676713	Hs.191155	ESTs	3.34	5.28
	312088	AW303760	Hs.13685	ESTs	1.60	1.15
	312093	T91809	Hs.121296	ESTs	0.68	0.85
75	312094	Z78390		gb:HSZ78390 Human fetal brain S. Meier-E	3.05	4.48
	312097	AI352096	Hs.112180	zinc finger protein 148 (pHZ-52)	4.52	9.70
	312118	T85332	Hs.178294	ESTs	2.40	2.60
	312128	AI052609	Hs.17631	Homo sapiens cDNA FLJ20118 fis, clone CO	2.39	3.53
	312147	T89855	Hs.195648	ESTs	0.67	1.03
80	312175	AA953383	Hs.127554	ESTs	5.85	10.60
	312179	AI052572	Hs.269864	ESTs	2.41	3.32
	312201	AI928365	Hs.91139	solute carrier family 1 (neuronal/epithe	0.24	0.89
	312207	H90213	Hs.191330	ESTs	2.20	4.55
	312220	N74613		gb:za55a07.s1 Soares fetal liver spleen	4.28	11.13
85	312252	AI128388	Hs.143655	ESTs	1.64	1.57
	312304	AA491949	Hs.269392	ESTs	0.12	2.47



	312318	AW235092	Hs.143981	ESTs	3.46	5.69
	312319	AA216698	Hs.180780	TERA protein	5.78	4.46
	312321	R66210	Hs.186937	ESTs	0.44	1.74
5	312331	AA825512	Hs.289101	glucose regulated protein, 58kD	3.73	5.96
	312339	AA524394	Hs.165544	ESTs	3.07	0.95
	312363	AI675558	Hs.181867	ESTs	10.08	16.73
	312375	AI375096	Hs.172405	cell division cycle 27	2.78	3.71
	312376	R52089	Hs.172717	ESTs	1.00	1.00
10	312389	AI863140		gb:tz43h12.x1 NCI_CGAP_Brn52 Homo sapien	2.37	3.98
	312437	AA995028		gb:RC4-BT0629-120200-011-b10 BT0629 Homo	4.06	5.41
	312440	AI051133	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	1.00	1.00
	312451	R59989	Hs.176539	ESTs	4.96	10.04
	312458	AI167637	Hs.146924	ESTs	1.11	1.00
15	312507	AI168177	Hs.143653	ESTs	5.89	8.24
	312520	AI742591	Hs.205392	ESTs	3.30	8.92
	312548	AI566228	Hs.159426	hypothetical protein PRO2121	1.38	1.65
	312564	H21520	Hs.35088	ESTs	0.40	0.77
	312583	AI193122	Hs.124141	ESTs	0.13	0.94
20	312599	AI865073	Hs.125720	ESTs	3.75	5.29
	312602	AA046451	Hs.165200	ESTs	6.78	12.93
	312645	H52121	Hs.193007	ESTs	0.38	1.13
	312666	AI240582	Hs.214678	ESTs	0.98	2.03
	312689	AW450461	Hs.203965	ESTs	0.21	0.61
25	312817	H75459	Hs.233425	ESTs	1.51	0.85
	312846	AW152104	Hs.200879	ESTs	8.93	13.78
	312873	AI690071	Hs.283552	ESTs, Weakly similar to unnamed protein	4.20	6.23
	312893	AI016204	Hs.172922	ESTs	2.67	3.15
	312902	AW292797	Hs.130316	ESTs, Weakly similar to T2D3_HUMAN TRANS	1.19	0.71
30	312925	N90868	Hs.271695	ESTs	2.50	4.25
	312936	AI681581	Hs.121525	ESTs	1.00	1.17
	312975	AI640506	Hs.293119	ESTs, Weakly similar to ALU7_HUMAN ALU S	2.30	4.80
	312978	N24887	Hs.292500	ESTs	0.80	1.05
	312980	AA497043	Hs.115685	ESTs	3.12	3.60
35	312984	N25871	Hs.177337	ESTs	2.03	2.13
	313000	AI147412	Hs.146657	ESTs	5.52	8.42
	313029	AA731520	Hs.170504	ESTs	0.96	1.39
	313039	AI419290	Hs.149990	ESTs, Weakly similar to unnamed protein	6.48	13.20
	313049	AW293055	Hs.119357	ESTs	6.44	10.73
40	313056	AI651930	Hs.135684	ESTs	1.51	2.04
	313058	D81015	Hs.125382	ESTs	0.25	1.50
	313070	AI422023	Hs.161338	ESTs	8.56	11.60
	313097	AI676164	Hs.204339	ESTs	3.72	4.56
	313130	AW449171	Hs.168677	ESTs	3.28	5.06
45	313136	N59284	Hs.288010	ESTs	0.49	1.36
	313153	AI240838	Hs.132750	ESTs	5.36	5.52
	313210	N74077	Hs.197043	ESTs	0.30	0.66
	313236	AW238169	Hs.83513	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.16	8.76
	313239	W19632	Hs.124170	ESTs	1.00	3.87
50	313265	N93466	Hs.121764	ESTs, Weakly similar to testicular tekti	0.74	2.06
	313267	AI770008	Hs.129583	ESTs	0.23	1.30
	313275	AI027604	Hs.159650	ESTs	6.68	9.57
	313290	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	1.34	1.07
	313292	AI362991	Hs.202121	ESTs, Weakly similar to env protein [H.s	2.00	4.32
55	313325	AI420611	Hs.127832	ESTs	1.20	2.27
	313357	AW074848	Hs.201501	ESTs	4.02	5.33
	313393	AI674685	Hs.200141	ESTs	1.36	2.84
	313399	AW376889	Hs.194097	ESTs	2.58	5.26
	313414	AI241540	Hs.132933	ESTs	6.57	15.07
60	313417	AA741151	Hs.137323	ESTs	0.63	3.01
	313457	AA576052	Hs.193223	Homo sapiens cDNA FLJ11646 fis, clone HE	2.78	4.70
	313499	AI261390	Hs.146085	KIAA1345 protein	0.91	2.37
	313516	AA029058	Hs.135145	ESTs	3.41	7.08
	313556	AA628517	Hs.118502	ESTs	0.23	0.70
65	313569	AI273419	Hs.135146	hypothetical protein FLJ13984	1.88	1.00
	313570	AA041455	Hs.209312	ESTs	0.73	2.27
	313638	AI753075	Hs.104627	Homo sapiens cDNA FLJ10158 fis, clone HE	1.00	1.72
	313662	AA740151	Hs.130425	ESTs	0.20	1.42
	313671	W49823	Hs.104613	RP42 homolog	1.00	1.00
70	313672	AW468891	Hs.122948	ESTs	3.46	5.80
	313690	AI493591	Hs.78146	platelet/endothelial cell adhesion molec	0.51	0.97
	313711	AA398070	Hs.133471	ESTs	0.18	1.01
	313723	AA070412		gb:zm68c10.s1 Stratagene neuroepithelium	1.08	1.03
	313726	AI744687	Hs.257806	ESTs	2.13	2.99
75	313774	AW136836	Hs.144583	ESTs	1.38	1.19
	313784	AA910514	Hs.134905	ESTs	3.88	5.78
	313790	AW078569	Hs.177043	ESTs	0.22	2.06
	313832	AW271022	Hs.133294	ESTs	1.15	0.91
	313834	AW418779	Hs.114889	ESTs	0.68	3.14
80	313835	AI538438	Hs.159087	ESTs	5.74	8.88
	313852	H18633	Hs.123641	protein tyrosine phosphatase, receptor t	0.16	1.14
	313854	AW470806	Hs.275002	ESTs	2.09	4.06
	313865	AA731470	Hs.163839	ESTs	3.41	4.09
	313871	AW471088	Hs.145950	ESTs	5.28	6.83
85	313883	AI949384		gb:nu76d01.s1 NCI_CGAP_Alvi1 Homo sapiens	2.90	10.91
	313915	AI969390	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	1.00	1.00



	313926	AW473830	Hs.171442	ESTs	3.40	4.11
	313948	AW452823	Hs.135268	ESTs	5.77	9.15
	313978	AI870175	Hs.13957	ESTs	0.46	0.75
5	313983	AI829133	Hs.226780	ESTs	4.10	6.40
	314035	AA164199	Hs.270152	ESTs	5.88	7.90
	314037	AW300048	Hs.275272	ESTs	1.00	3.79
	314040	AA166970	Hs.118748	ESTs	7.60	11.33
	314067	AW293538	Hs.51743	KIAA1340 protein	1.86	1.21
10	314103	AI028477	Hs.132775	ESTs	2.90	5.29
	314107	AA806113	Hs.189025	ESTs	2.00	1.66
	314113	AA218986	Hs.118854	ESTs	0.91	4.17
	314124	AW118745	Hs.9460	Homo sapiens mRNA; cDNA DKFZp547C244 (fr	2.53	3.32
	314126	AA226431		gb:nc18b12.s1 NCI_CGAP_Pr1 Homo sapiens	3.13	5.08
15	314128	AA935633	Hs.194628	ESTs	2.90	6.35
	314151	AA236163	Hs.202430	ESTs	4.15	6.45
	314184	AW081795	Hs.233465	ESTs	3.44	4.65
	314192	AW290975	Hs.118923	ESTs	1.00	1.23
	314244	AL036450	Hs.103238	ESTs	2.88	3.67
20	314253	AA278679	Hs.189510	ESTs	4.98	7.16
	314262	AW086215	Hs.246096	ESTs	0.38	1.94
	314320	AA811598	Hs.275809	ESTs	3.34	5.66
	314332	AL037551	Hs.95612	ESTs	2.85	2.09
	314335	AA287443	Hs.142570	Homo sapiens clone 24629 mRNA sequence	4.35	4.78
25	314340	AW304350	Hs.130879	ESTs, Moderately similar to putative p15	0.77	0.86
	314351	AA292275	Hs.193746	ESTs	3.07	3.77
	314376	AI628633	Hs.324679	ESTs	4.10	6.11
	314443	AA827125	Hs.192043	ESTs	6.20	13.67
	314458	AI217440	Hs.143873	ESTs	0.58	2.49
30	314466	AA767818	Hs.122707	ESTs	2.53	2.62
	314478	AI521173	Hs.125507	DEAD-box protein	3.94	5.65
	314482	AL043807	Hs.134182	ESTs	1.30	1.44
	314506	AA833655	Hs.206868	Homo sapiens cDNA FLJ14056 fis, clone HE	3.28	3.47
	314519	R42554	Hs.210862	T-box, brain, 1	3.12	6.16
35	314529	AL046412	Hs.202151	ESTs	3.43	6.87
	314546	AW007211	Hs.16131	hypothetical protein FLJ12876	1.38	1.00
	314562	AI564127	Hs.143493	ESTs	2.29	5.27
	314579	AW197442	Hs.116998	ESTs	3.87	5.75
	314580	AW451832	Hs.255938	ESTs, Moderately similar to KIAA1200 pro	0.10	0.71
40	314585	AA918474	Hs.216363	ESTs	1.08	1.40
	314589	AW384790	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	1.00	1.00
	314592	AA435761	Hs.192148	ESTs	0.90	2.60
	314603	AA418024	Hs.270670	ESTs	4.56	6.29
	314604	AA946582	Hs.8700	deleted in liver cancer 1	3.42	3.92
45	314606	AA418241	Hs.188767	ESTs	2.97	4.55
	314648	AA878419		gb:EST391378 MAGE resequences, MAGP Homo1.42		1.36
	314699	AI038719	Hs.132801	ESTs	3.66	4.97
	314701	AI754634	Hs.131987	ESTs	0.03	0.90
	314710	AI669131	Hs.290989	EST	3.40	7.52
50	314750	AI095005	Hs.135174	ESTs	2.80	6.54
	314767	AW135412	Hs.164002	ESTs	3.20	4.26
	314801	AA481027	Hs.109045	hypothetical protein FLJ10498	1.00	1.00
	314817	AI694139	Hs.192855	ESTs	0.91	0.99
	314835	AI281370	Hs.76064	ribosomal protein L27a	5.75	7.44
55	314852	AI903735		gb:MR-BT035-200199-031 BT035 Homo sapien	1.68	4.34
	314853	AA729232	Hs.153279	ESTs	0.60	1.85
	314940	AW452768	Hs.162045	ESTs	10.10	16.20
	314941	AA515902	Hs.130650	ESTs	0.31	1.02
	314943	AI476797	Hs.184572	cell division cycle 2, G1 to S and G2 to	2.18	0.37
60	314955	AA521382	Hs.192534	ESTs	2.59	3.90
	314973	AW273128	Hs.300268	ESTs	1.05	1.25
	315004	AA527941	Hs.325351	EST	5.64	13.63
	315006	AI538613	Hs.298241	Transmembrane protease, serine 3	0.52	1.78
65	315033	AI493046	Hs.146133	ESTs	2.46	1.00
	315035	AI569476	Hs.177135	ESTs	0.34	1.33
	315056	AI202703	Hs.152414	ESTs	2.10	2.64
	315069	AI821517	Hs.105866	ESTs	1.00	1.30
	315071	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 fis, clone C	1.78	1.00
70	315073	AW452948	Hs.257631	ESTs	1.17	1.52
	315078	AA568548	Hs.190616	ESTs	3.00	3.79
	315080	AA744550	Hs.136345	ESTs	1.00	1.00
	315120	AA564991	Hs.269477	ESTs	0.64	1.44
	315175	AI025842	Hs.152530	ESTs	0.61	1.91
	315193	AI241331	Hs.131765	ESTs	1.06	0.97
75	315196	AA972756	Hs.44898	Homo sapiens clone TCCCTA00151 mRNA sequ	0.48	1.96
	315200	AI808235	Hs.307686	EST	3.76	9.40
	315254	AI474433	Hs.179556	ESTs	5.37	9.36
	315353	AW452608	Hs.279610	hypothetical protein FLJ10493	1.00	1.30
	315397	AA218940	Hs.137516	fidgetin-like 1	3.38	2.24
80	315403	AW362980	Hs.163924	ESTs	2.04	5.23
	315431	AA622104	Hs.184838	ESTs	2.36	8.04
	315454	AI239473		gb:qh36f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.46	7.64
	315455	AW393391	Hs.156919	ESTs	3.78	5.76
	315473	AI681671	Hs.312671	ESTs, Moderately similar to OVCA1	0.89	2.15
	315483	AW512763	Hs.222024	transcription factor BMAL2	2.32	1.96
85	315526	AI193048	Hs.128685	ESTs	1.67	1.78



	315530	AI200852	Hs.127780	ESTs	1.05	1.01
	315541	AI168233	Hs.123159	sperm associated antigen 4	0.85	0.56
	315552	AW445034	Hs.256578	ESTs	1.00	2.22
5	315562	AA737415	Hs.152826	ESTs	2.66	2.48
	315577	AW513545	Hs.17283	hypothetical protein FLJ10890	2.20	2.25
	315587	AI268399	Hs.140489	ESTs	1.00	1.04
	315589	AW072387	Hs.158258	Homo sapiens mRNA; cDNA DKFZp434B1272 (f	0.14	1.05
	315623	AA364078	Hs.258189	ESTs	7.44	12.56
10	315634	AA837085	Hs.220585	ESTs	0.50	1.40
	315668	AA912347	Hs.136585	ESTs	0.43	1.22
	315677	AI932662	Hs.164073	ESTs	0.60	1.39
	315706	AW440742	Hs.155556	hypothetical protein FLJ20202	2.18	3.77
	315707	AI418055	Hs.161160	ESTs	2.88	2.63
15	315730	H25899	Hs.201591	ESTs	0.11	0.60
	315745	AI821759	Hs.191856	ESTs	3.50	7.25
	315791	AA678177		gb:zi15a05.s1 Soares_fetal_liver_spleen_	1.78	2.63
	315801	AA827752	Hs.266134	ESTs	4.31	6.23
	315820	AI652022	Hs.258785	ESTs	2.35	3.01
20	315878	AA683336	Hs.189046	ESTs	2.12	2.64
	315905	AI821911	Hs.209452	ESTs	1.03	1.97
	315923	AI052789	Hs.133263	ESTs	2.63	5.06
	315954	AW276810	Hs.254859	ESTs, Moderately similar to ALU5_HUMAN A	1.21	0.85
	315978	AA830893	Hs.119769	ESTs	3.09	3.41
25	316001	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	2.20	6.82
	316011	AW516953	Hs.201372	ESTs	0.35	1.63
	316012	AA764950	Hs.119898	ESTs	6.56	8.13
	316040	AI983409	Hs.189226	ESTs	5.69	10.69
	316048	AI720759	Hs.224971	ESTs	2.84	10.45
30	316076	AW297895	Hs.116424	ESTs	0.30	1.05
	316124	AI308862	Hs.167028	ESTs	1.00	1.43
	316151	AI806016	Hs.156520	ESTs	5.80	9.03
	316187	AW518299	Hs.192253	ESTs	1.20	3.96
	316204	AA731509	Hs.120257	ESTs	4.92	6.94
35	316232	AW297853	Hs.251203	ESTs	1.48	1.60
	316275	AI671041	Hs.292611	ESTs, Moderately similar to ALU1_HUMAN A	5.86	12.14
	316291	AW375974	Hs.156704	ESTs	2.73	2.69
	316303	AA740994	Hs.209609	ESTs	1.53	1.26
	316344	AA744518	Hs.120610	ESTs	3.66	8.34
40	316346	AI028478	Hs.157447	ESTs	3.51	6.69
	316365	AI627845	Hs.210776	ESTs	2.50	4.33
	316380	AI393378	Hs.164496	ESTs	1.16	2.16
	316470	AA809902	Hs.243813	ESTs	5.40	10.34
	316509	AA767310	Hs.291766	ESTs	2.46	2.89
45	316514	AA768037	Hs.291671	ESTs	4.70	6.04
	316519	AI929097		gb:od10c11.s1 NCI_CGAP_GCB1 Homo sapiens	4.41	9.70
	316609	AW292520	Hs.122082	ESTs	1.00	2.89
	316633	AI125586	Hs.127955	ESTs	2.61	3.72
	316700	AW172316	Hs.252961	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.46	4.64
50	316711	AI743721	Hs.285316	ESTs, Moderately similar to ALU7_HUMAN A	4.45	6.95
	316713	AI090671	Hs.134807	hypothetical protein FLJ12057	0.30	2.40
	316715	AI440266	Hs.170673	ESTs, Weakly similar to AF126780 1 retin	0.20	1.45
	316787	AW369770	Hs.130351	ESTs	4.05	5.53
	316809	AA825839	Hs.202238	ESTs	2.25	3.82
55	316811	AA922060	Hs.132471	ESTs	1.00	1.32
	316812	AW135045	Hs.232001	ESTs	3.28	4.70
	316818	AA827176	Hs.124316	ESTs	0.67	1.81
	316824	AA837416	Hs.124299	ESTs	3.53	6.00
	316827	AI380429	Hs.172445	ESTs	0.72	1.56
60	316891	AW298119	Hs.202536	ESTs	1.64	2.97
	316951	AA134365	Hs.57548	ESTs	1.45	1.08
	316970	AA860172	Hs.132406	ESTs	1.00	1.53
	316971	AA860212	Hs.170991	ESTs	1.08	1.96
	316990	AA861611	Hs.130643	ESTs	5.44	10.04
65	317001	AI627917	Hs.233694	hypothetical protein FLJ11350	3.56	4.37
	317008	AW051597	Hs.143707	ESTs	0.69	1.37
	317051	AA873253	Hs.126233	ESTs	6.18	12.72
	317128	AA971374	Hs.125674	ESTs	1.87	2.66
	317129	H12523	Hs.78521	Homo sapiens cDNA: FLJ21193 fis, clone C	4.12	6.64
70	317137	AW341567	Hs.125710	ESTs	2.82	5.12
	317196	AI348258	Hs.153412	ESTs	1.98	2.51
	317212	AI866468	Hs.148294	ESTs	1.86	2.83
	317223	AW297920	Hs.130054	ESTs	0.83	1.57
	317224	D56760	Hs.93029	sparc/osteonectin, cwcv and kazal-like d	2.74	0.86
75	317266	AA906289	Hs.203614	ESTs	1.00	1.00
	317282	AI807444	Hs.176101	ESTs	2.60	4.21
	317285	AW370882	Hs.222080	ESTs	1.96	3.49
	317302	AA908709	Hs.135564	ESTs	7.16	8.32
	317304	AW449899	Hs.130184	ESTs	1.38	2.28
80	317320	AA927151	Hs.130452	ESTs	3.58	8.13
	317413	AW341701	Hs.126622	ESTs	2.08	4.92
	317417	AA918420	Hs.145378	ESTs	3.06	4.79
	317452	AA972965	Hs.135568	ESTs	4.22	9.21
	317519	AI859695	Hs.126860	ESTs	1.88	4.15
85	317521	AI824338	Hs.126891	ESTs	3.12	4.55
	317529	AI916517	Hs.126865	ESTs	2.73	3.34



	317570	AI733361	Hs.127122	ESTs	1.00	2.43
	317571	AA938663	Hs.199828	ESTs	5.20	11.95
	317598	AW206035	Hs.192123	ESTs	0.33	1.56
5	317627	AI346110	Hs.132553	ESTs	1.50	1.39
	317650	AI733310	Hs.127346	ESTs	0.48	1.46
	317659	AA961216	Hs.127785	ESTs	4.18	7.14
	317674	AW294909	Hs.132208	ESTs	2.92	3.20
	317686	AA969051	Hs.187319	ESTs	1.00	1.01
10	317692	AI307659	Hs.174794	ESTs	5.33	9.59
	317701	AI674774	Hs.128014	ESTs	1.00	1.00
	317711	AI733015	Hs.272189	ESTs	5.13	7.81
	317722	AI733373	Hs.128119	ESTs	2.50	6.03
	317756	AA973667	Hs.128320	ESTs	1.59	1.30
15	317777	AI143525	Hs.47313	KIAA0258 gene product	1.00	2.48
	317799	AI498273	Hs.128808	ESTs	1.78	2.11
	317803	AA983251	Hs.128899	ESTs	0.80	1.06
	317821	AI368158	Hs.70983	PTPL1-associated RhoGAP 1	0.17	0.68
	317848	AI820575	Hs.129086	Homo sapiens cDNA FLJ12007 fis, clone HE	5.30	8.16
20	317850	N29974	Hs.152982	hypothetical protein FLJ13117	1.30	2.28
	317861	AW341064	Hs.129119	ESTs	2.18	5.93
	317865	AI298794	Hs.129130	ESTs	4.48	8.20
	317869	AW295184	Hs.129142	deoxyribonuclease II beta	0.44	0.99
	317881	AI827248	Hs.224398	Homo sapiens cDNA FLJ11469 fis, clone HE	4.06	2.23
25	317890	AI915599	Hs.129225	ESTs	4.68	7.48
	317899	AI952430	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	3.14	3.37
	317986	AI005163	Hs.201378	ESTs, Weakly similar to T12545 hypotheti	0.28	1.66
	318001	AW235697	Hs.130980	ESTs	5.12	9.97
	318016	AI016694	Hs.256921	ESTs	1.86	4.50
30	318023	AW243058	Hs.131155	ESTs	2.92	5.22
	318054	AW449270	Hs.232140	ESTs	3.92	6.37
	318068	AI024540	Hs.131574	ESTs	1.21	1.27
	318117	AI208304	Hs.250114	ESTs	0.86	1.17
	318187	AI792585	Hs.133272	ESTs, Weakly similar to ALUC_HUMAN !!!!	5.90	6.98
35	318223	AI077540	Hs.134090	ESTs	1.05	0.90
	318240	AI085377	Hs.143610	ESTs	3.10	2.40
	318255	AI082692	Hs.134662	ESTs	0.02	1.05
	318266	AI554341	Hs.271443	ESTs	6.12	10.55
	318330	AI093840	Hs.143758	ESTs	4.98	7.90
40	318369	AI493501	Hs.170974	ESTs	2.46	5.62
	318428	AI949409	Hs.194591	ESTs	0.77	0.45
	318458	AI149783	Hs.158438	ESTs	3.54	4.92
	318467	AI151395	Hs.144834	ESTs	4.56	5.62
45	318473	AI939339	Hs.146883	ESTs	2.08	4.05
	318476	AI693927	Hs.265165	ESTs	4.22	8.07
	318487	AI167877	Hs.143716	ESTs	1.47	1.05
	318488	AI217431	Hs.144709	ESTs	1.40	4.14
	318491	T26477	Hs.22883	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.84	1.90
	318499	T25451		gb:PTHI188 HTCDL1 Homo sapiens cDNA 5/3	2.58	5.20
50	318537	AA377908	Hs.13254	ESTs	3.26	4.18
	318538	N28625	Hs.74034	Homo sapiens clone 24651 mRNA sequence	0.35	1.07
	318547	R20578	Hs.90431	ESTs	3.22	4.60
	318552	R18364	Hs.90363	ESTs	4.87	9.06
	318575	R55102	Hs.107761	ESTs, Weakly similar to unnamed protein	1.91	1.98
55	318580	T34571	Hs.49007	poly(A) polymerase alpha	2.74	6.22
	318587	AA779704	Hs.168830	Homo sapiens cDNA FLJ12136 fis, clone MA	0.85	2.46
	318596	AI470235	Hs.172698	EST	4.88	4.93
	318622	T48325	Hs.237658	apolipoprotein A-II	4.80	12.51
	318629	N25163	Hs.8861	ESTs	0.39	1.04
60	318637	AA243539	Hs.9196	hypothetical protein	1.72	3.57
	318648	T77141	Hs.184411	albumin	6.27	9.91
	318650	AA393302	Hs.176626	hypothetical protein EDAG-1	3.96	8.84
	318671	AA188823	Hs.299254	Homo sapiens cDNA: FLJ23597 fis, clone L	1.53	0.81
	318679	T58115	Hs.10336	ESTs	1.00	2.19
65	318711	AI936475	Hs.101282	Homo sapiens cDNA: FLJ21238 fis, clone C	3.05	3.18
	318725	AI962487	Hs.242990	ESTs	1.08	2.46
	318728	Z30201	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S	0.77	1.33
	318740	NM_002543	Hs.77729	oxidised low density lipoprotein (lectin	0.25	1.49
	318776	R24963	Hs.23766	ESTs	1.00	3.01
70	318784	H00148	Hs.5181	proliferation-associated 2G4, 38kD	2.70	3.86
	318816	F07873	Hs.21273	ESTs	3.90	7.13
	318865	H10818		gb:ym04f10.r1 Soares infant brain 1NIB H	2.25	3.56
	318879	R56332	Hs.18268	adenylate kinase 5	1.78	5.00
	318881	Z43224	Hs.124952	ESTs	4.79	14.13
75	318894	F08138	Hs.7387	DKFZP564B116 protein	5.31	7.00
	318901	AW368520	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	1.03	0.91
	318925	Z43577	Hs.21470	ESTs	2.23	3.80
	318936	AI219221	Hs.308298	ESTs	1.86	7.16
	318982	Z44140	Hs.269622	ESTs	5.84	9.79
80	318986	Z44186	Hs.169161	ESTs, Highly similar to MAON_HUMAN NADP-	1.00	1.00
	319041	Z44720	Hs.98365	ESTs, Weakly similar to weak similarity	3.38	6.11
	319103	H05896	Hs.4993	KIAA1313 protein	1.00	1.07
	319170	R13678	Hs.285306	putative selenocysteine lyase	3.79	5.03
	319196	F07953	Hs.16085	putative G-protein coupled receptor	1.00	2.98
85	319199	F07361	Hs.13306	ESTs	3.53	5.66
	319242	F11472	Hs.12839	ESTs	5.87	7.26



	319263	T65331	Hs.81360	Homo sapiens cDNA: FLJ21927 fis, clone H	1.81	1.57
	319267	F11802	Hs.6818	ESTs	1.10	4.72
	319270	R13474	Hs.290263	ESTs	4.80	10.40
5	319279	T65094	Hs.12677	CGI-147 protein	1.50	2.11
	319282	AA461358	Hs.12876	ESTs	1.00	1.00
	319289	W07304	Hs.79059	transforming growth factor, beta recepto	0.18	0.68
	319291	W86578	Hs.285243	hypothetical protein FLJ22029	0.26	0.62
	319293	F12119	Hs.12583	ESTs	3.13	4.50
10	319312	Z45481		gb:HSC2QE041 normalized infant brain cDN	1.10	1.00
	319370	H54254	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	0.16	0.73
	319391	R06304	Hs.13911	ESTs	1.26	2.43
	319396	H67130	Hs.301743	ESTs	0.70	0.76
	319398	AA359754	Hs.191196	ESTs	2.45	3.59
15	319407	R05329		gb:ye91b04.r1 Soares fetal liver spleen	2.00	3.54
	319425	T82930		gb:yd39f07.r1 Soares fetal liver spleen	4.28	8.81
	319433	R06050	Hs.191198	ESTs	6.15	14.13
	319437	AA282420	Hs.111991	ESTs, Weakly similar to Y48A5A.1 [C.eleg	3.26	5.68
	319466	AI809937	Hs.116417	ESTs	1.76	5.65
20	319471	R06546	Hs.19717	ESTs	4.29	4.84
	319480	R06933	Hs.184221	ESTs	1.00	1.00
	319484	T91772		gb:yd52a10.s1 Soares fetal liver spleen	2.81	4.88
	319486	AI382429	Hs.250799	ESTs	2.08	2.82
	319508	T99898	Hs.270104	ESTs, Moderately similar to ALU8_HUMAN A	2.80	4.39
25	319523	T69499	Hs.191184	ESTs	1.55	3.25
	319545	R83716	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	1.65	1.19
	319546	R09692		gb:yf23b12.r1 Soares fetal liver spleen	5.11	8.54
	319552	AA096106	Hs.20403	ESTs	1.89	3.36
	319582	T82998	Hs.250154	hypothetical protein FLJ12973	3.48	4.82
30	319586	D78808	Hs.283683	chromosome 8 open reading frame 4	0.26	0.82
	319604	R11679	Hs.297753	vimentin	1.68	3.41
	319609	AW247514	Hs.12293	hypothetical protein FLJ21103	3.06	4.24
	319611	H14957		gb:ym19c10.r1 Soares infant brain 1NIB H	2.76	4.24
	319653	AA770183	Hs.173515	uncharacterized hypothalamus protein HT0	2.51	3.55
35	319657	R19897	Hs.106604	ESTs	5.32	7.68
	319658	R13432	Hs.167481	syntrophin, gamma 1	3.35	5.00
	319661	H08035	Hs.21398	ESTs, Moderately similar to A Chain A, H	5.18	12.55
	319662	H06382	Hs.21400	ESTs	1.58	1.56
	319708	R15372	Hs.22664	ESTs	1.00	1.22
40	319742	T77668	Hs.21162	ESTs	2.48	3.13
	319748	R18178	Hs.295866	Homo sapiens mRNA; cDNA DKFZp434N1923 (f	3.02	4.85
	319772	R76633	Hs.22646	ESTs	4.36	11.61
	319788	AA321932	Hs.117414	KIAA1320 protein	2.56	3.68
	319805	R92857	Hs.271350	likely ortholog of mouse polydom	4.63	6.56
45	319812	N74880	Hs.264330	N-acylsphingosine amidohydrolase (acid c	0.63	1.32
	319834	AA071267		gb:zm61g01.r1 Stratagene fibroblast (937	0.30	0.94
	319878	T78517	Hs.13941	ESTs	3.99	6.44
	319882	AA258981	Hs.291392	ESTs	5.09	7.36
	319912	T77559	Hs.94109	Homo sapiens cDNA FLJ13634 fis, clone PL	3.24	3.21
50	319935	H79460	Hs.271722	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.40	9.42
	319944	T79248	Hs.133510	ESTs	3.31	5.39
	319947	AA160967	Hs.14479	Homo sapiens cDNA FLJ14199 fis, clone NT	2.90	4.95
	319962	H06350	Hs.135056	Human DNA sequence from clone RP5-850E9	1.81	1.57
55	320007	AA336314		gb:EST40943 Endometrial tumor Homo sapie	3.42	6.29
	320018	T83263		gb:yd40h09.r1 Soares fetal liver spleen	2.77	5.14
	320030	H63789	Hs.296288	ESTs, Weakly similar to KIAA0638 protein	4.10	6.69
	320032	AI699772	Hs.292664	ESTs, Weakly similar to A46010 X-linked	3.27	3.27
	320040	AA233671	Hs.87164	hypothetical protein FLJ14001	1.81	1.64
	320047	T86564	Hs.302256	EST	3.38	7.36
60	320063	AA074108	Hs.120844	FOXJ2 forkhead factor	5.90	16.73
	320096	H58138	Hs.117915	ESTs	2.08	4.47
	320099	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	1.00	1.00
	320112	T92107	Hs.188489	ESTs	2.27	2.06
	320140	H94179	Hs.119023	SMC2 (structural maintenance of chromoso	1.00	1.00
65	320188	AW419200	Hs.172318	ESTs	1.26	1.00
	320193	AA831259	Hs.17132	ESTs	2.58	6.23
	320195	R62203	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	2.85	4.53
	320199	R78659	Hs.29792	ESTs	0.40	0.94
	320203	AL049227	Hs.124776	Homo sapiens mRNA; cDNA DKFZp564N1116 (f	0.84	1.18
70	320219	AA327564	Hs.127011	tubulointerstitial nephritis antigen	1.00	1.17
	320220	AF054910	Hs.127111	tektin 2 (testicular)	0.18	1.09
	320225	AF058989	Hs.128231	G antigen, family B, 1 (prostate associa	5.26	13.75
	320231	H03139	Hs.24683	ESTs	1.59	1.93
	320260	NM_003608	Hs.131924	G protein-coupled receptor 65	1.38	4.56
75	320267	AL049337	Hs.132571	Homo sapiens mRNA; cDNA DKFZp564P016 (fr	1.00	1.92
	320268	H06019	Hs.151293	Homo sapiens cDNA FLJ10664 fis, clone NT	5.58	5.70
	320322	AF077374	Hs.139322	small proline-rich protein 3	1.41	1.01
	320325	AI167978	Hs.139851	caveolin 2	0.05	0.67
	320330	AF026004	Hs.141660	chloride channel 2	2.17	1.26
80	320339	H10807	Hs.281434	Homo sapiens cDNA FLJ14028 fis, clone HE	1.81	2.32
	320388	H16065	Hs.31286	ESTs	1.00	3.22
	320402	R22291	Hs.23368	Homo sapiens clone FLC0578 PRO2852 mRNA,	1.41	1.36
	320413	AA203711	Hs.173269	ESTs	2.31	3.61
	320432	R62786	Hs.124136	ESTs	11.25	20.78
85	320436	AA253352	Hs.293663	ESTs	2.22	3.49
	320438	W24548	Hs.5669	ESTs	3.53	8.14



	320448	AI240233	Hs.80887	v-yes-1 Yamaguchi sarcoma viral related	1.42	3.46
	320451	R26944	Hs.180777	Homo sapiens mRNA; cDNA DKFZp564M0264 (f	0.87	0.81
	320484	AA094436	Hs.296267	folliculin-like 1	0.65	1.18
5	320499	R32555	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	3.44	7.15
	320514	AB007978	Hs.158278	KIAA0509 protein	6.44	13.62
	320521	N31464	Hs.24743	hypothetical protein FLJ20171	1.48	1.04
	320526	AW374205	Hs.111314	ESTs	3.66	7.87
	320527	R34672	Hs.324522	ESTs	3.16	5.63
10	320536	AA331732	Hs.137224	ESTs	2.83	5.83
	320556	AF054177	Hs.14570	hypothetical protein FLJ22530	1.28	1.00
	320564	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.22	0.81
	320587	Z44524	Hs.167456	Homo sapiens mRNA full length insert cDN	1.84	2.44
	320635	R54159	Hs.80506	small nuclear ribonucleoprotein polypept	1.00	6.25
15	320639	AA243258	Hs.7395	hypothetical protein FLJ23182	2.60	2.30
	320648	N48521	Hs.26549	Homo sapiens mRNA for KIAA1708 protein,	1.00	1.53
	320651	AA489268	Hs.111334	ferritin, light polypeptide	0.14	0.79
	320664	AI904216	Hs.91251	hypothetical protein FLJ11198	5.02	8.84
	320676	AA132650	Hs.300511	ESTs	3.63	5.37
20	320683	R59291	Hs.26638	ESTs, Weakly similar to unnamed protein	0.37	1.31
	320689	AA334609	Hs.171929	ESTs, Weakly similar to A54849 collagen	1.27	1.02
	320696	AW135016	Hs.172780	ESTs	3.53	4.60
	320714	AI445591		gb:yq04a10.r1 Soares fetal liver spleen	1.06	0.85
	320727	U96044	Hs.181125	immunoglobulin lambda locus	1.35	1.49
25	320771	AI793266	Hs.117176	poly(A)-binding protein, nuclear 1	0.04	0.82
	320794	AA281993	Hs.91226	ESTs	2.96	4.33
	320822	AF100780	Hs.194679	WNT1 inducible signaling pathway protein	0.10	0.79
	320824	AF120274	Hs.194689	artemin	1.16	1.11
	320830	AJ132445	Hs.266416	claudin 14	1.06	1.75
30	320843	AA317372	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	1.36	1.47
	320849	D60031	Hs.34771	ESTs	5.30	7.49
	320853	AI473796	Hs.135904	ESTs	1.00	1.00
	320896	AB002155	Hs.271580	uroplakin 1B	5.90	2.55
	320921	R94038	Hs.199538	inhibin, beta C	2.20	1.17
35	320927	AI205786	Hs.213923	ESTs	0.18	1.46
	320957	AI878933	Hs.92023	core histone macroH2A2.2	1.67	2.18
	320997	H22544		gb:yn69f11.r1 Soares adult brain N2b5HB5	3.26	3.62
	321045	W88483	Hs.293650	ESTs	2.25	4.55
	321046	H27794	Hs.269055	ESTs	2.69	4.25
40	321052	AW372884	Hs.240770	nuclear cap binding protein subunit 2, 2	2.14	2.56
	321059	AI092824	Hs.126465	ESTs	1.69	0.53
	321062	R87955	Hs.241411	Homo sapiens mRNA full length insert cDN	2.76	5.20
	321067	AF131782	Hs.241438	Homo sapiens clone 24941 mRNA sequence	4.79	7.41
	321102	AA018306		gb:ze40d08.r1 Soares retina N2b4HR Homo	1.79	4.27
45	321130	H43750	Hs.125494	ESTs	1.00	3.14
	321142	AI817933	Hs.298351	ASPL protein	8.73	15.36
	321155	AA336635	Hs.99598	hypothetical protein MGC5338	3.04	5.03
	321158	AA700289		gb:yu76f11.r1 Soares fetal liver spleen	4.62	8.39
	321170	N53742	Hs.172982	ESTs	2.21	4.46
50	321199	AW385512		gb:yy56d10.s1 Soares_multiple_sclerosis_	5.69	8.01
	321206	H54178	Hs.226469	Homo sapiens cDNA FLJ12417 fis, clone MA	4.00	7.32
	321225	AL080073	Hs.251414	Homo sapiens mRNA; cDNA DKFZp564B1462 (f	4.17	4.63
	321236	AW371941	Hs.18192	Ser/Arg-related nuclear matrix protein (	1.00	1.00
	321244	AF068654		gb:Homo sapiens isolate AN.1 immunoglobu	2.18	9.13
55	321270	R83560		gb:yv76c06.s1 Soares fetal liver spleen	3.80	5.26
	321317	AI937060	Hs.6298	KIAA1151 protein	1.81	1.65
	321318	AB033041	Hs.137507	KIAA1215 protein	1.00	1.00
	321325	AB033100	Hs.300646	KIAA protein (similar to mouse paladin)	0.44	0.93
	321342	AA127984	Hs.222024	transcription factor BMAL2	4.94	4.93
60	321356	R93443	Hs.271770	ESTs	3.10	4.66
	321418	AI739161	Hs.161075	ESTs	2.28	2.54
	321420	AI368667	Hs.132743	ESTs	1.13	0.97
	321430	U05890		gb:H.sapiens (DIG3) mRNA for immunoglobu	2.42	3.35
	321453	N50080	Hs.82845	Homo sapiens cDNA: FLJ21930 fis, clone H	1.60	3.11
65	321467	X13075		gb:Human 2a12 mRNA for kappa-immunoglobu	0.42	0.72
	321468	AA514198	Hs.38540	ESTs	2.46	6.50
	321491	H70665	Hs.292549	ESTs	1.00	1.25
	321498	AW295517	Hs.255436	ESTs	3.19	6.24
	321504	W02356	Hs.268980	ESTs	2.28	3.86
70	321510	AA703650	Hs.255748	ESTs	2.14	3.94
	321513	H84972	Hs.108551	ESTs	2.78	5.37
	321516	AI382803	Hs.159235	ESTs	3.06	7.19
	321565	AI525773	Hs.266514	hypothetical protein FLJ11342	4.89	7.82
	321577	H84260		gb:ys90g04.r1 Soares retina N2b5HR Homo	1.00	1.73
75	321581	AA019964	Hs.28803	ESTs	4.88	6.73
	321582	AA143755	Hs.21858	trinucleotide repeat containing 3	1.00	2.08
	321587	H95531		gb:ys76e02.r1 Soares retina N2b4HR Homo	2.26	4.52
	321626	AA295430	Hs.96322	hypothetical protein FLJ23560	1.95	3.83
	321628	H87064	Hs.161051	ESTs, Moderately similar to ALU6_HUMAN A	0.47	1.02
80	321642	AW085917	Hs.247084	ESTs	1.52	1.38
	321669	H95404	Hs.294110	ESTs	2.17	2.45
	321687	AA625149		gb:af70c12.r1 Soares_NhHMPu_S1 Homo sapi	4.31	6.95
	321688	H97646	Hs.123158	Homo sapiens cDNA FLJ12830 fis, clone NT	2.82	3.28
	321693	AA700017	Hs.173737	ras-related C3 botulinum toxin substrate	0.51	1.08
85	321700	N55160	Hs.167260	ESTs	4.57	7.46
	321701	AW390923	Hs.42568	ESTs	1.00	1.00



	321709	N25847	Hs.108923	RAB38, member RAS oncogene family	1.00	1.00
	321710	N35682	Hs.259743	ESTs	2.97	5.26
	321775	AI694875	Hs.202312	Homo sapiens clone N11 NTera2D1 teratoca	1.00	1.00
	321777	AI637993	Hs.202312	Homo sapiens clone N11 NTera2D1 teratoca	1.68	0.45
5	321779	N42729	Hs.163835	ESTs	0.90	0.90
	321829	D81993	Hs.8966	tumor endothelial marker 8	2.69	3.89
	321846	AA281594	Hs.87902	ESTs	5.11	7.64
	321879	AL109670	Hs.302809	ESTs	6.49	9.58
10	321883	AA426494	Hs.46901	KIAA1462 protein	0.28	0.95
	321899	N55158	Hs.29468	ESTs	0.39	0.95
	321911	AF026944	Hs.293797	ESTs	6.20	10.76
	321949	R49202	Hs.181694	EST	4.62	10.51
	321955	AI651866	Hs.195689	ESTs	2.89	5.47
15	321956	AL110177	Hs.132882	ESTs	0.32	1.25
	321987	AL133612	Hs.272759	KIAA1457 protein	1.00	1.83
	321991	AL133627	Hs.158923	Homo sapiens mRNA; cDNA DKFZp434K0722 (f	4.00	6.47
	322002	AA328801	Hs.84522	ESTs	2.10	3.48
	322035	AL137517	Hs.306201	hypothetical protein DKFZp564O1278	1.00	1.90
20	322044	AW340926		gb:xy51b10.x1 NCI_CGAP_Lu34.1 Homo sapie	3.20	9.67
	322057	N92197	Hs.154679	synaptotagmin 1	1.55	1.07
	322060	AI341937		gb:qt10e03.x1 NCI_CGAP_GC4 Homo sapiens	4.59	7.68
	322070	U80769	Hs.210322	Homo sapiens mRNA for KIAA1766 protein,	2.78	4.52
	322083	AF074982	Hs.226031	ESTs, Highly similar to KIAA0535 protein	3.10	5.52
	322091	AI819863	Hs.106243	ESTs	1.59	1.75
25	322125	R93901		gb:yq16c12.r1 Soares fetal liver spleen	2.06	5.27
	322130	R98978	Hs.117767	ESTs	10.12	16.49
	322147	AF085919	Hs.114176	ESTs	0.94	0.64
	322166	AF085958		gb:yr88b03.r1 Soares fetal liver spleen	4.09	6.67
30	322173	H52567		gb:yt85d04.r1 Soares_pineal_gland_N3HPG	3.46	4.85
	322178	H56535		gb:yt88g03.r1 Soares_pineal_gland_N3HPG	0.44	2.54
	322179	H92891		gb:yt94c02.s1 Soares_pineal_gland_N3HPG	4.52	7.50
	322186	H67346	Hs.269187	ESTs	0.15	0.98
	322196	W87895	Hs.211516	ESTs	2.20	5.04
35	322212	AF087995	Hs.134877	ESTs	3.42	4.84
	322221	AI890619	Hs.179662	nucleosome assembly protein 1-like 1	0.82	2.14
	322277	AI640193	Hs.226389	ESTs	3.62	3.98
	322278	AF086283		gb:zd46f01.r1 Soares_fetal_heart_NbHH19W	1.00	1.00
	322284	AI792140	Hs.49265	ESTs	0.66	2.76
40	322288	AL037273	Hs.7886	pellino (Drosophila) homolog 1	0.71	0.70
	322320	AF086419		gb:zd78d03.r1 Soares_fetal_heart_NbHH19W	2.02	2.76
	322336	AA308526	Hs.76152	decorin	2.92	4.44
	322339	W17348		gb:zb18c07.x5 Soares_fetal_lung_NbHL19W	8.50	11.56
	322366	AW404274	Hs.122492	hypothetical protein	0.61	1.34
45	322372	W25624	Hs.153943	ESTs	7.37	12.07
	322374	AI394663	Hs.122116	ESTs, Moderately similar to Osf2 [M.musc	4.78	10.50
	322378	AF064819	Hs.201877	DESC1 protein	1.00	1.00
	322388	AI815730	Hs.247474	hypothetical protein FLJ21032	7.09	8.49
	322416	AA223183	Hs.298442	adaptor-related protein complex 3, mu 1	3.20	5.80
50	322419	AA248987	Hs.14084	ring finger protein 7	1.64	1.57
	322425	W37943	Hs.34892	KIAA1323 protein	0.83	1.00
	322431	AA069222	Hs.141892	ESTs	3.96	5.22
	322450	AA040131	Hs.25144	ESTs	5.18	12.67
	322465	AA137152	Hs.286049	phosphoserine aminotransferase	3.41	2.23
55	322467	AF116826	Hs.180340	putative protein-tyrosine kinase	1.00	1.30
	322473	AA744286	Hs.266935	tRNA selenocysteine associated protein	1.75	2.03
	322509	T52172	Hs.302213	ESTs	1.00	2.27
	322523	W80398	Hs.193197	ESTs	2.75	5.49
	322527	AF147359		gb:Homo sapiens full length insert cDNA	1.25	1.27
60	322560	AI916847	Hs.270947	ESTs	4.57	8.81
	322566	W87285	Hs.269587	ESTs	1.00	1.42
	322585	AA837622		gb:zh69c01.r1 Soares_fetal_liver_spleen_	4.18	6.94
	322635	AA679084		gb:zh90h08.r1 Soares_fetal_liver_spleen_	2.40	4.85
	322641	AA007352	Hs.256042	ESTs	2.94	4.64
65	322653	AI828854	Hs.258538	striatin, calmodulin-binding protein	0.48	0.38
	322664	AA011522		gb:zi03g07.r1 Soares_fetal_liver_spleen_	1.92	2.18
	322687	AI110759		gb:AF074666 Human fetal liver cDNA libra	4.14	6.75
	322692	AA018117	Hs.60843	potassium voltage-gated channel, shaker-	3.50	5.00
	322694	AI110872	Hs.279812	PRO0327 protein	1.80	1.72
70	322708	AF113674	Hs.283773	clone FLB1727	1.00	3.43
	322712	AA021328	Hs.23607	hypothetical protein FLJ11109	3.28	3.86
	322766	AW068805	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.63	1.53
	322770	AA045796	Hs.122682	ESTs	1.53	1.06
	322794	AI608591	Hs.38991	S100 calcium-binding protein A2	12.06	1.94
75	322810	AI962276	Hs.127444	ESTs	4.09	6.90
	322818	AW043782	Hs.293616	ESTs	1.20	1.63
	322820	AI377755	Hs.120695	ESTs	0.21	1.93
	322872	AA827228	Hs.126943	ESTs	2.04	1.63
	322882	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	5.26	1.22
80	322887	AI986306	Hs.86149	phosphoinositol 3-phosphate-binding prot	2.80	2.24
	322913	AI733737	Hs.68837	ESTs	2.38	6.61
	322926	AI825940	Hs.211192	ESTs	4.02	5.79
	322929	AI365585	Hs.146246	ESTs	0.30	1.14
	322968	AI905228	Hs.83484	SRV (sex determining region Y)-box 4	2.06	1.13
85	322971	C15953	Hs.212760	hypothetical protein FLJ13649	1.18	2.00
	322981	AA493252	Hs.159577	ESTs	2.28	2.61



	322988	C18727	Hs.171941	ESTs	0.39	2.00
	323003	AI733859	Hs.149089	ESTs	3.28	1.00
	323013	AA134042	Hs.191451	ESTs	3.38	5.68
5	323025	AL157565	Hs.315369	Homo sapiens cDNA: FLJ23075 fis, clone L	0.06	1.10
	323032	AW244073	Hs.145946	ESTs	10.18	21.27
	323052	R21124	Hs.85573	Homo sapiens DC29 mRNA, complete cds	1.46	1.90
	323064	AL119341	Hs.49359	Homo sapiens mRNA; cDNA DKFZp547E052 (fr	3.08	5.64
	323098	AI700025	Hs.270471	ESTs	2.31	4.49
10	323102	AL119913	Hs.163615	ESTs	5.38	11.64
	323155	AL135041		gb:DKFZp762K2310_r1 762 (synonym: hmel2)	2.38	5.56
	323176	AW071648	Hs.82101	pleckstrin homology-like domain, family	1.06	1.41
	323191	AA195600	Hs.301570	ESTs	0.73	1.24
	323225	AA205654	Hs.24790	KIAA1573 protein	5.25	11.95
15	323232	AA148722	Hs.224680	ESTs	0.45	1.35
	323266	AW003362	Hs.243886	nuclear autoantigenic sperm protein (his	1.71	1.83
	323281	AI697556	Hs.292659	ESTs	1.24	3.21
	323283	AA256014	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	12.68	15.05
	323314	AA226310	Hs.191501	ESTs	4.42	9.61
20	323316	AL134620	Hs.280175	ESTs	2.98	5.93
	323334	AI336501	Hs.77273	ras homolog gene family, member A	1.98	3.30
	323338	R74219	Hs.23348	S-phase kinase-associated protein 2 (p45	1.62	1.00
	323348	AA233056	Hs.191518	ESTs	1.00	1.07
	323351	AA704103	Hs.24049	ESTs	1.43	1.68
25	323359	AA234172	Hs.137418	ESTs	0.34	1.18
	323360	AA716061	Hs.161719	ESTs	3.01	3.71
	323405	AW139550	Hs.115173	ESTs	1.90	8.81
	323420	AI672386	Hs.263780	ESTs	0.29	1.01
	323434	AW081455	Hs.120219	ESTs	2.27	1.92
30	323445	AA253103	Hs.135569	ESTs, Weakly similar to NEUROD [H.sapien	0.43	0.80
	323449	AA282865	Hs.284153	Fanconi anemia, complementation group A	3.19	3.85
	323492	H00978	Hs.20887	hypothetical protein FLJ10392	2.70	3.20
	323501	AA182461	Hs.84520	ESTs	2.04	3.31
	323505	AI652287		gb:EST382593 MAGE resequences, MAGK Homo2.21		3.08
35	323515	AA282274	Hs.256083	ESTs	2.69	3.40
	323541	AI185116	Hs.104613	RP42 homolog	1.20	1.09
	323545	AI814405	Hs.224569	ESTs	1.25	1.55
	323635	R63117	Hs.9691	Homo sapiens cDNA: FLJ23249 fis, clone C	0.27	0.72
	323675	AA984759	Hs.272168	tumor differentially expressed 1	3.70	5.80
40	323678	AL042121	Hs.20880	ESTs	3.33	5.10
	323691	AA317561	Hs.145599	ESTs	1.00	1.00
	323693	AW297758	Hs.249721	ESTs	2.01	1.54
	323746	AW298611	Hs.12808	MARK	4.11	5.53
	323774	AA329806	Hs.321056	Homo sapiens mRNA; cDNA DKFZp586F1322 (f	2.06	3.70
45	323856	AA355264	Hs.267604	hypothetical protein FLJ10450	3.42	8.13
	323857	T18988	Hs.293668	ESTs	5.97	12.51
	323870	AA341774	Hs.129212	ESTs	3.17	4.52
	323876	AL042492	Hs.147313	ESTs	0.36	1.00
	323885	AA344308	Hs.128427	Homo sapiens BAC clone RP11-335J18 from	2.31	3.33
50	323911	AL043212	Hs.92550	ESTs	4.38	5.41
	323919	AA862973	Hs.220704	ESTs	5.80	10.20
	323972	AI869964	Hs.182906	ESTs	3.10	5.14
	324005	AA610011	Hs.208021	ESTs	5.34	10.07
	324036	AI472078	Hs.303662	ESTs	1.00	5.03
55	324055	AA528794	Hs.128644	ESTs	0.86	1.00
	324063	AW292740	Hs.272813	dual oxidase 1	0.45	0.91
	324072	AA381829		gb:EST94855 Activated T-cells I Homo sap	2.82	5.12
	324092	AW269931	Hs.202473	Homo sapiens cDNA: FLJ22278 fis, clone H	2.40	2.52
	324095	AW377983	Hs.298140	Homo sapiens cDNA: FLJ22502 fis, clone H	1.32	4.30
60	324129	AI381918	Hs.285833	Homo sapiens cDNA: FLJ22135 fis, clone H	1.40	1.77
	324132	AW504860	Hs.288836	hypothetical protein FLJ12673	4.24	6.21
	324214	AA412395	Hs.225740	ESTs	6.96	10.69
	324227	AA295552	Hs.28631	Homo sapiens cDNA: FLJ22141 fis, clone H	0.81	0.53
	324266	AL047634	Hs.231913	ESTs	2.42	4.05
65	324275	AA429088	Hs.98523	ESTs	3.62	5.38
	324281	AL048026	Hs.124675	ESTs, Weakly similar to T14742 hypotheti	0.14	0.70
	324290	AA432032	Hs.304420	ESTs	3.71	4.34
	324303	AL118754		gb:DKFZp761P1910_r1 761 (synonym: hamy2)	0.95	0.91
	324312	AI198841	Hs.128173	ESTs	4.06	5.91
70	324325	AL138153	Hs.300410	ESTs	5.88	8.25
	324338	AL138357	Hs.145078	regulator of differentiation (in S. pomb	0.87	1.25
	324341	AW197734	Hs.99807	ESTs, Weakly similar to unnamed protein	1.28	1.00
	324343	AW452016	Hs.293232	ESTs	2.54	3.46
	324371	AA452305	Hs.270319	ESTs	5.85	8.36
75	324382	AW502749	Hs.24724	MFH-amplified sequences with leucine-ric	0.76	1.64
	324384	AA453396	Hs.127656	KIAA1349 protein	2.88	5.69
	324385	F28212	Hs.284247	KIAA1491 protein	1.81	1.99
	324388	AI924963	Hs.306206	hypothetical protein FLJ11215	1.00	1.00
	324432	AA464510	Hs.152812	ESTs	2.73	2.17
	324497	AW152624	Hs.136340	ESTs, Weakly similar to unnamed protein	0.71	1.90
80	324510	AI148353	Hs.287425	Homo sapiens cDNA FLJ11569 fis, clone HE	1.00	1.00
	324580	AA492588		gb:ng99c08.s1 NCI_CGAP_Thy1 Homo sapiens	2.18	3.50
	324582	AA506935	Hs.132036	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.96	11.36
	324633	AA572994	Hs.325489	ESTs	2.92	4.22
	324640	AW295832	Hs.134798	ESTs, Moderately similar to TTL MOUSE TU	5.48	11.74
85	324675	AW014734	Hs.157969	ESTs	0.39	0.73



	324699	AW504732	Hs.21275	hypothetical protein FLJ11011	0.93	0.93
	324747	AA603532	Hs.130807	ESTs	1.57	1.81
	324748	AA657457	Hs.292385	ESTs	1.55	1.34
5	324801	AI819924	Hs.14553	sterol O-acyltransferase (acyl-Coenzyme	1.00	6.56
	324804	AI692552		gb:wd73f12.x1 NCI_CGAP_Lu24 Homo sapiens	1.00	7.53
	324828	AA843926	Hs.124434	ESTs	2.00	3.25
	324855	AW152305	Hs.122364	ESTs	2.74	3.43
	324866	AI541214	Hs.46320	Small proline-rich protein SPRK [human,	1.07	0.95
10	324871	AW297755	Hs.271923	Homo sapiens cDNA: FLJ22785 fis, clone K	1.68	1.21
	324886	AA806794	Hs.131511	ESTs	2.56	5.61
	324889	D31010		gb:HUML12147 Human fetal lung Homo sapie	2.20	4.65
	324948	AW383618	Hs.265459	ESTs, Moderately similar to ALU2_HUMAN A	5.28	7.05
	324953	AI264628	Hs.125428	ESTs	3.37	5.51
15	324958	AA625076	Hs.132892	protocadherin 20	5.12	9.81
	324988	T06997	Hs.121028	hypothetical protein FLJ10549	2.52	1.08
	325024	F13254	Hs.78672	laminin, alpha 4	5.24	10.22
	325105	H97109	Hs.105421	ESTs	1.00	1.00
	325108	AA401863	Hs.22380	ESTs	1.99	2.14
20	325114	D83901	Hs.315562	ESTs	2.73	3.17
	325146	AI064690	Hs.171176	ESTs	1.86	3.41
	325149	D61117	Hs.187646	ESTs	0.42	0.93
	325187	AI653682	Hs.197812	ESTs	6.50	11.31
	325228				6.18	15.76
25	325235				2.64	4.12
	325328				2.87	4.42
	325340				0.29	0.33
	325367				16.56	24.29
	325373				0.63	1.22
30	325389				0.88	1.05
	325436				5.75	14.14
	325471				8.46	17.82
	325498				3.32	6.42
	325557				5.51	8.28
35	325559				7.48	21.40
	325560				4.08	6.25
	325569				4.20	5.24
	325585				1.10	1.13
	325587				1.00	1.00
40	325597				2.98	13.40
	325639				0.78	0.78
	325685				0.46	0.66
	325686				0.95	1.55
	325735				4.48	9.20
45	325739				0.59	0.88
	325740				2.42	6.61
	325792				7.88	9.83
	325819				4.74	7.18
	325883				2.02	2.64
50	325895				7.78	15.98
	325925				2.04	10.60
	325932				4.18	7.36
	325941				3.66	9.03
	325969				0.61	0.80
55	325971				4.88	7.42
	326025				0.55	1.07
	326046				7.21	14.72
	326099				3.60	5.98
	326108				1.27	1.06
60	326163				3.27	5.70
	326165				0.45	1.11
	326189				0.13	0.45
	326204				5.60	9.00
	326230				7.00	12.01
65	326274				1.00	8.09
	326360				9.86	15.35
	326393				0.52	0.77
	326505				1.00	1.42
	326515				1.24	5.84
70	326589				9.20	13.49
	326592				2.77	4.01
	326605				2.01	2.53
	326692				1.00	1.00
	326693				1.00	1.31
75	326720				0.19	0.65
	326742				2.34	7.20
	326770				0.25	0.83
	326818				3.09	4.56
	326936				2.08	3.45
80	326964				0.41	1.70
	326983				2.02	3.80
	326991				1.09	1.20
	327036				1.00	8.04
	327040				3.05	4.22
85	327053				3.55	6.31
	327075				1.59	1.40



	327085	2.50	12.57
	327130	5.38	8.04
	327156	3.74	6.58
5	327220	1.28	1.54
	327224	6.56	12.91
	327288	2.61	5.40
	327321	2.42	3.11
	327332	6.62	10.58
10	327361	2.69	4.41
	327377	2.04	6.72
	327396	2.61	4.50
	327414	1.00	8.01
	327442	5.91	9.65
15	327467	6.58	18.01
	327473	3.79	7.48
	327483	4.08	8.87
	327562	0.68	2.86
	327568	1.00	2.00
20	327606	2.06	3.61
	327611	5.90	14.26
	327642	4.06	8.74
	327654	1.05	2.08
	327734	1.00	1.00
25	327775	1.46	11.79
	327796	3.47	5.65
	327840	3.26	6.64
	327940	5.84	15.58
	327984	0.36	1.50
30	328004	1.87	1.42
	328021	0.42	0.59
	328068	2.83	4.68
	328100	3.04	5.39
	328101	3.54	5.20
35	328113	0.72	0.91
	328157	5.58	5.16
	328196	5.76	11.13
	328197	5.98	10.58
	328264	3.11	4.88
40	328299	2.20	3.06
	328342	1.49	1.94
	328365	1.00	1.00
	328369	4.40	7.36
	328381	1.86	4.93
45	328451	5.51	7.56
	328481	0.13	0.72
	328500	2.71	3.97
	328530	5.41	7.62
	328600	3.14	10.68
50	328608	4.56	8.17
	328616	2.24	11.91
	328623	3.04	5.46
	328632	0.70	1.19
	328664	3.48	6.80
55	328666	10.42	26.47
	328698	9.68	14.56
	328700	2.74	10.22
	328708	0.15	0.57
	328735	6.23	8.91
60	328743	3.62	6.54
	328806	0.22	0.78
	328861	3.68	10.54
	328908	5.42	16.36
	328933	2.02	5.29
65	328934	1.73	4.45
	328949	3.34	5.41
	329005	2.88	7.26
	329011	2.52	3.72
	329033	1.00	1.03
70	329037	5.07	8.16
	329067	1.98	2.41
	329134	2.24	3.25
	329157	2.30	11.04
	329178	2.64	5.02
75	329192	6.41	15.27
	329194	0.31	0.79
	329204	1.60	3.75
	329224	2.99	6.11
	329228	0.83	0.83
80	329288	0.63	1.01
	329337	1.00	1.00
	329541	0.76	1.68
	329560	1.34	2.02
	329588	1.68	2.22
85	329643	4.18	11.77
	329703	1.00	1.00



	329764			5.78	15.50
	329816			2.09	5.44
	329860			3.13	10.77
	329993			7.83	14.21
5	330020			5.58	13.12
	330036			3.32	5.57
	330052			4.31	7.97
	330085			1.34	1.76
	330088			4.70	12.46
10	330093			0.44	1.06
	330100			3.47	4.83
	330106			2.14	3.61
	330107			3.17	6.87
	330120			5.61	11.89
15	330123			4.50	12.74
	330208			1.55	7.62
	330263			13.10	23.38
	330300			2.81	4.98
	330313			3.00	4.41
20	330366			0.67	0.76
	330372			4.76	11.82
	330385	AA449749	Hs.182971	2.14	2.15
	330397	D14659	Hs.154387	0.40	1.15
25	330468	L10343	Hs.112341	1.11	0.94
	330472	L24203	Hs.82237	1.67	1.17
	330478	L38486	Hs.296049	0.46	1.07
	330493	M27826	Hs.267319	1.07	0.95
	330495	M31328	Hs.71642	0.97	0.96
	330506	M61906	Hs.6241	0.17	3.66
30	330512	M80563	Hs.81256	0.60	1.06
	330537	U19765	Hs.2110	2.81	2.07
	330547	U32989	Hs.183671	3.91	1.49
	330551	U39840	Hs.299867	1.15	1.03
35	330568	U56244		2.83	4.79
	330599	U90437		2.08	1.54
	330601	U90916	Hs.82845	0.89	1.35
	330605	X02419	Hs.77274	1.87	1.55
	330609	X04741	Hs.76118	1.83	1.30
40	330617	X53587	Hs.85266	1.54	1.15
	330630	X78669	Hs.79088	1.39	1.19
	330644	Y07755	Hs.38991	3.83	1.13
	330650	Z68228	Hs.2340	1.25	0.95
	330660	AA347868	Hs.139293	15.50	29.07
45	330692	AA017045	Hs.6702	1.00	1.00
	330707	AA133891	Hs.293690	0.20	1.35
	330715	AA233707	Hs.11571	0.12	1.40
	330717	AA233926	Hs.52620	6.62	5.42
	330722	AA243560	Hs.34382	1.40	1.65
50	330740	AA297746	Hs.22654	0.27	2.04
	330742	AA400979	Hs.25691	0.44	0.90
	330744	AA406142	Hs.12393	0.71	3.23
	330751	AA428286	Hs.29643	1.66	1.52
	330760	AA448663	Hs.30469	0.52	0.90
55	330763	AA450200	Hs.274337	0.37	0.97
	330786	D60374	Hs.49136	0.78	0.84
	330790	T48536	Hs.105807	0.23	3.17
	330814	AA015730	Hs.265398	0.37	2.07
	330827	AA040332	Hs.12744	1.60	1.00
60	330844	AA063037	Hs.66803	0.93	1.16
	330901	AA157818	Hs.267319	1.02	1.03
	330931	F01443	Hs.284256	0.24	0.88
	330952	H02855	Hs.29567	0.08	1.31
	330961	H10998	Hs.7164	1.29	1.26
65	330968	H16568	Hs.23748	0.48	0.96
	331014	H98597	Hs.30340	0.29	0.74
	331046	N66563	Hs.191358	0.99	8.56
	331060	N75081	Hs.157148	1.24	1.00
	331099	R36671	Hs.83937	0.75	1.03
70	331108	R41408	Hs.21983	1.00	2.75
	331131	R54797		6.04	10.68
	331135	R61398	Hs.4197	0.80	0.96
	331170	T23461	Hs.159293	2.63	4.29
	331180	T32446	Hs.6640	1.78	2.71
75	331183	T40769	Hs.8469	1.00	3.01
	331203	T82310		1.70	3.80
	331271	AA059347	Hs.82226	1.20	3.19
	331306	AA252079	Hs.63931	0.31	1.30
	331327	AA281076	Hs.109221	2.09	2.41
80	331341	AA303125	Hs.23240	0.72	2.43
	331359	AA416979	Hs.46901	0.09	0.91
	331363	AA421562	Hs.91011	1.02	0.87
	331378	AA448881	Hs.49282	1.03	1.23
	331384	AA456001	Hs.93847	1.40	1.00
	331402	AA505135	Hs.44037	1.80	3.93
85	331422	F10802	Hs.163628	1.65	1.89



	331490	N32912	Hs.26813	CDA14	2.48	1.73
	331531	N51343		gb:yz15g04.s1 Soares_multiple_sclerosis_	0.98	1.68
	331547	N54811		gb:od74f04.s1 NCI_CGAP_Ov2 Homo sapiens	3.80	5.75
5	331578	N67960	Hs.249989	ESTs	0.11	0.67
	331589	N71027	Hs.152618	ESTs	1.09	1.38
	331608	N89861	Hs.112110	PTD007 protein	0.93	0.76
	331614	N92293	Hs.240272	EST	0.17	1.34
	331668	W69707	Hs.58030	EST	2.24	3.82
10	331671	W72033	Hs.194695	ras homolog gene family, member I	1.00	1.24
	331676	W79834	Hs.58559	ESTs, Weakly similar to rhotekin [M.musc	0.08	1.07
	331681	W85712	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	8.72	4.27
	331692	W93592	Hs.152213	wingless-type MMTV integration site fami	0.94	0.54
	331717	AA190888	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par	1.57	1.34
15	331718	AA191404	Hs.104072	ESTs	6.80	11.77
	331811	AA404500	Hs.301570	ESTs	1.10	1.00
	331820	AA405970	Hs.97996	transcription termination factor, mitoc	0.73	0.59
	331831	AA412031	Hs.97901	EST	2.77	4.08
	331852	AA418988	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	0.23	0.93
20	331943	AA453418	Hs.21275	hypothetical protein FLJ11011	0.36	1.88
	331969	AA460702	Hs.82772	collagen, type XI, alpha 1	1.00	1.00
	331990	AA478102	Hs.139631	ESTs	3.04	3.87
	332002	AA482009	Hs.105104	ESTs	1.19	0.78
	332027	AA489671	Hs.65641	hypothetical protein FLJ20073	1.27	1.03
25	332029	AA489697	Hs.145053	ESTs	0.30	1.62
	332033	AA489840	Hs.251014	EST	2.30	3.70
	332048	AA496019	Hs.201591	ESTs	0.17	0.52
	332071	AA598594	Hs.205293	KIAA1211 protein	1.35	1.23
	332074	AA599012		gb:ae41e11.s1 Gessler Wilms tumor Homo s	0.19	2.00
30	332083	AA600200	Hs.155546	KIAA1080 protein; Golgi-associated, gamm	0.31	1.18
	332085	AA600353	Hs.173933	nuclear factor I/A	0.30	1.50
	332125	AA609861	Hs.312447	ESTs	0.22	0.62
	332177	F10812	Hs.101433	ESTs	8.21	18.03
	332180	H03348	Hs.7327	claudin 1	2.27	1.57
35	332185	H10356	Hs.101689	ESTs	0.09	1.18
	332203	H49388	Hs.317769	EST	8.05	5.02
	332232	N48891	Hs.101915	Stargardt disease 3 (autosomal dominant)	0.78	0.85
	332240	N54803	Hs.324267	ESTs, Weakly similar to putative p150 [	0.96	1.23
	332261	N70294	Hs.269137	ESTs	2.40	3.74
40	332275	R08838	Hs.26530	serum deprivation response (phosphatidyl	0.27	0.75
	332280	R38100	Hs.146381	RNA binding motif protein, X chromosome	0.39	1.88
	332299	R69250	Hs.21201	necln 3; DKFZP566B0846 protein	5.24	12.76
	332304	R74041	Hs.101539	ESTs	1.44	3.18
	332314	T25862	Hs.101774	hypothetical protein FLJ23045	0.68	1.32
45	332384	M11433	Hs.101850	retinol-binding protein 1, cellular	1.71	0.88
	332434	N75542	Hs.289068	Homo sapiens cDNA FLJ11918 fis, clone HE	0.43	0.86
	332445	T63781	Hs.11112	ESTs	0.68	1.00
	332453	L00205	Hs.111758	keratin 6A	31.54	1.00
	332458	M33493	Hs.250700	tryptase beta 1	0.51	1.00
50	332504	AA053917	Hs.15106	chromosome 14 open reading frame 1	0.79	1.24
	332525	M17252	Hs.278430	cytochrome P450, subfamily XXIA (steroid	0.98	1.70
	332530	M31682	Hs.1735	inhibin, beta B (activin AB beta polypep	0.88	0.66
	332535	N20284	Hs.19280	cysteine-rich motor neuron 1	0.22	1.46
	332539	AA412528	Hs.20183	ESTs, Weakly similar to AF164793 1 prote	0.93	1.49
55	332559	M13955	Hs.166189	cytokeratin 2	0.35	1.13
	332563	N92924	Hs.274407	protease, serine, 16 (thymus)	1.00	1.00
	332565	AA234896	Hs.25272	E1A binding protein p300	0.36	1.05
	332594	AA279313	Hs.3239	methyl CpG binding protein 2 (Rett syndr	0.53	0.59
	332634	S38953	Hs.283750	tenascin XA	0.38	1.16
60	332638	AA283034	Hs.50640	JAK binding protein	1.00	1.70
	332640	AA417152	Hs.5101	protein regulator of cytokinesis 1	6.15	1.16
	332654	AA001296	Hs.288217	hypothetical protein MGC2941	1.50	2.73
	332665	AA223335	Hs.63788	propionyl Coenzyme A carboxylase, beta p	1.20	0.91
	332692	AA496035	Hs.247926	gap junction protein, alpha 5, 40kD (con	0.17	1.12
65	332716	L00058	Hs.79070	v-myc avian myelocytomatosis viral oncog	1.00	1.44
	332736	L13773	Hs.114765	myeloid/lymphoid or mixed-lineage leukem	1.00	1.81
	332758	X93921	Hs.296938	dual specificity phosphatase 7	0.53	0.78
	332781	AA233258	Hs.247112	hypothetical protein FLJ10902	1.44	1.56
	332792				1.70	1.19
70	332816				1.85	2.47
	332858				1.04	1.57
	332906				3.48	8.04
	332911				1.00	1.00
	332912				1.06	4.40
75	332922				1.00	1.00
	332956				0.42	0.88
	332959				1.96	6.34
	332982				0.56	0.99
	332984				0.30	0.78
80	332998				1.47	2.01
	333058				0.47	1.38
	333097				2.14	3.19
	333121				2.76	3.70
	333122				1.92	1.21
85	333123				1.85	1.39
	333138				0.47	0.52



	333139	1.88	0.84
	333140	0.21	0.64
	333221	1.51	1.11
	333260	0.75	1.01
5	333380	6.68	15.75
	333387	4.56	12.61
	333512	5.05	8.01
	333524	2.28	3.98
10	333585	2.31	1.53
	333603	2.23	1.17
	333604	2.51	1.58
	333618	0.52	0.98
	333627	1.44	1.36
15	333628	1.90	1.90
	333650	1.85	2.10
	333678	1.85	2.35
	333750	2.18	5.67
	333763	1.99	2.60
20	333767	1.02	0.96
	333768	1.78	1.65
	333769	2.15	2.13
	333772	1.46	2.53
	333777	1.00	1.42
25	333846	2.99	4.50
	333884	0.47	0.94
	333887	0.50	1.00
	333891	0.43	0.89
	333892	0.51	0.91
30	333904	0.26	1.13
	333906	0.55	0.98
	333948	1.70	2.15
	333954	0.37	1.09
	333966	8.10	14.30
35	333968	0.63	1.38
	334061	4.24	12.30
	334094	1.30	12.03
	334113	4.55	8.63
	334161	0.82	1.59
40	334183	0.47	0.76
	334187	1.36	3.70
	334219	0.69	1.04
	334222	1.88	1.70
	334223	4.72	3.14
45	334239	0.79	0.62
	334255	0.45	1.10
	334333	1.00	3.56
	334378	3.98	5.76
	334382	1.50	1.31
50	334492	3.59	4.75
	334562	5.94	15.40
	334588	8.14	19.53
	334616	1.55	1.56
	334633	5.16	8.07
55	334648	0.59	2.13
	334787	3.70	7.15
	334866	8.13	10.60
	334891	0.32	1.14
	334933	1.00	3.84
60	334934	4.01	7.43
	334945	1.04	2.96
	334967	0.29	1.14
	334990	1.50	1.39
	335015	5.88	18.65
65	335093	0.55	1.75
	335120	4.31	8.01
	335125	0.38	1.97
	335179	1.24	1.98
	335188	0.46	1.47
70	335211	1.61	1.42
	335288	0.73	0.97
	335289	0.20	0.26
	335361	2.18	1.58
	335379	0.50	0.71
75	335414	3.64	14.94
	335416	2.93	3.98
	335496	0.96	0.91
	335497	1.71	1.92
	335548	1.15	2.40
80	335551	3.22	10.54
	335558	3.42	4.89
	335586	5.50	12.75
	335619	2.99	3.07
	335620	3.80	8.29
85	335621	0.28	0.57
	335682	0.46	1.17



	335686	2.55	3.81
	335755	2.24	1.07
	335784	0.20	0.97
5	335814	1.13	1.48
	335815	2.45	3.51
	335823	1.00	4.16
	335835	0.49	1.70
	335851	1.66	1.39
10	335868	2.98	6.43
	335896	0.98	0.99
	335936	12.10	21.93
	335948	1.00	1.64
	335983	1.00	4.21
15	335995	0.37	1.17
	336021	1.04	0.84
	336034	11.40	23.54
	336038	1.19	1.21
	336066	0.54	1.63
20	336107	0.95	0.70
	336205	3.13	6.29
	336275	3.20	10.10
	336292	2.34	3.09
	336331	1.00	1.00
25	336419	0.65	0.79
	336632	2.33	2.16
	336633	2.55	2.23
	336634	2.19	2.03
	336635	2.69	2.48
30	336636	2.13	1.83
	336637	2.43	2.24
	336638	2.31	2.03
	336659	0.60	1.31
	336675	0.31	1.18
35	336684	1.50	1.14
	336694	4.74	7.10
	336716	4.43	6.37
	336721	2.20	0.74
	336798	1.64	2.14
40	336900	6.14	12.73
	336948	1.00	1.00
	337028	1.30	2.09
	337043	4.01	11.53
	337046	1.67	1.84
45	337054	2.78	7.35
	337128	7.20	16.14
	337162	3.45	5.34
	337183	5.72	11.41
	337184	3.72	5.90
50	337192	1.27	1.06
	337194	1.88	1.68
	337229	0.22	1.03
	337268	1.00	3.31
	337299	3.23	5.14
55	337325	2.76	3.72
	337389	5.80	10.42
	337493	2.06	6.30
	337497	7.88	20.29
	337500	3.80	4.48
60	337549	1.66	2.31
	337603	1.27	8.54
	337605	5.76	7.16
	337671	0.73	0.97
	337755	1.54	0.92
65	337786	5.07	9.73
	337809	6.18	12.87
	337862	3.78	12.97
	337871	2.66	8.16
	337958	0.26	1.34
70	338008	1.48	1.12
	338033	2.38	14.59
	338083	0.65	2.16
	338110	1.00	1.61
	338112	5.86	8.25
75	338145	1.70	1.97
	338148	8.07	18.19
	338158	1.30	4.55
	338161	2.58	3.57
	338179	1.00	1.00
80	338182	3.32	4.63
	338189	1.00	3.34
	338197	0.99	1.69
	338199	4.58	7.62
	338215	6.01	15.85
85	338279	0.53	0.95
	338316	20.58	38.66



	338322	3.23	7.39
	338357	4.10	11.39
	338359	10.12	21.59
5	338366	0.69	1.02
	338374	0.40	1.18
	338414	0.47	1.06
	338418	6.12	13.86
	338469	3.09	5.11
10	338501	6.28	10.32
	338506	6.97	12.41
	338523	3.10	5.84
	338549	1.70	2.70
	338561	0.79	0.81
15	338662	1.72	1.46
	338671	0.17	0.91
	338676	2.10	15.86
	338726	1.20	1.09
	338779	0.12	0.57
20	338804	0.99	1.67
	338836	1.00	1.00
	338871	4.30	9.81
	338872	5.02	12.81
	338879	0.23	1.12
25	338937	6.55	12.26
	338966	1.76	5.42
	338993	1.00	2.40
	339047	5.26	10.81
	339100	5.10	6.88
30	339114	1.00	1.70
	339121	1.00	3.75
	339170	10.36	19.67
	339229	4.08	13.48
	339264	2.64	3.83
35	339293	1.73	1.94

TABLE 8B shows the accession numbers for those Pkeys in Table 8A lacking unigenelD's. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

	Pkey:	Unique Eos probeset identifier number	
	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
45	Pkey	CAT number	Accessions
	322044	187363_1	AW340926 AA249063 N86075
	322060	44320_1	AI341937 AW003063 U34725 AA904742
50	321430	42705_1	X57414 X57415
	321467	43034_1	X13075 X13076
	322125	46779_1	R93901 AF075073 R93902
	322166	46861_1	H69434 AF085958 H69846
55	322173	46873_1	H52567 H52557 AF085970 H52164
	322178	46882_1	H56535 AF085980 H56712
	322179	46885_1	H92891 AF085982 H92777
	321577	1615102_1	H84849 H84252 H84260 H86664 H85320
	321587	1615333_1	H95531 H95521 H84529
60	313723	111953_1	AA070412 AA102346 AA081885
	320997	627492_1	H22544 H46842 AI204929
	322278	47271_1	W69304 AF086283 W69200
	321687	218439_1	AA625149 AA313030 AA313052 H97463
	313883	129439_1	AA665089 AA135130 AA484059 AA102419 AW877765
65	322320	47422_1	W79150 AF086419
	322339	814584_1	AI668646 AI734214 W17348
	314648	293660_1	AW979268 AA878419 AA431342 AA431628
	300201	682222_1	AI308300 AI308296
	306897	25196_-2	AI093967
70	323155	979809_1	AL120701 AL135041 AL121524
	322527	38927_1	AF147359 T58511 T58560
	322585	473768_2	W88919 W89125
	300362	1574395_1	Z42308 H23514
	322635	82296_1	AA005129 AA679084 AA694399
75	322664	85042_1	AA011522 AA702841 AA011691 AA330797
	315454	380580_1	AI239464 AI239473 AA625812 AI208703
	322687	37372_1	AF074666 AI110759 AF090902
	314852	327472_1	AI903735 AA491283 AI694953 AW976903 AA761362
	307783	697809_1	AI347274 AW844024
80	324072	269032_1	AA381722 AA381829 AW963906 AW963902 AA381242
	300627	221345_1	AA488472 W27363 AA317053 BE082689 AW967036 BE079872
	323505	196389_1	AW970512 AA280251 AI652287 BE466438 AI650725 AA551854 AA281574 AW571481
	315791	403558_1	AA678177 AA677034
	324303	233842_1	AL118754 AA333202 H38001
85	316519	442885_1	AA847835 AA768376
	300926	333127_1	AA504860 AA504911



	324580	328264_1	AA492588 AA492498 AA492571
	301882	275087_1	T78054 T79888 AA398185
	324804	398093_1	AI692552 AI393343 AI800510 AI377711 F24263 AA661876
5	324889	1515978_1	D31010 D30991 D31168 D31166 D31465
	302697	43219_1	AJ001409 AJ001410
	302711	45419_1	L08442 D51348
	302742	458_39	L12061
	318499	364430_1	T25451 AA585296 AA585305
10	310624	34624_4	U88896 U88898 AA916056 T03285 AI341594 AI359534 AI634031 U88897
	302847	458_105	X98941 X98942 X98943 X98953 X98949
	304122	77271_-5	H28966
	303598	270283_1	AA382814 AA402411 AA412355
	311409	837264_1	AI698839 AI909260 AI909259
15	312094	797889_1	Z78390 T97427
	319312	1540116_1	Z45481 F12393 T74437
	319407	1688823_1	R05329 R01555 R08276
	319425	1689571_1	T82930 R02424 T85145
	320007	229683_1	AA336314 T82938 AA327744 AW967388 AA639967 T10753
20	320018	1815987_1	T83263 T85731 T85730
	319484	1691553_1	T91772 R07257 R07098
	318865	1535937_1	H10818 F07831 Z43072
	312220	1671607_1	N74613 T98756 T98589
	319546	243305_1	R09692 R09414 AA346353
25	312389	902067_1	AI863140 W80703 R43474
	319611	1566863_1	H14957 R56522 R11908
	312437	291472_1	BE080180 AW827313 AW231970 AA995028 AA428584 AW872716 AW892508 AW854593 AA578441 AW975234 AA664937 AA984131 AA528743 AA552874 AA564758 AW063245 AI267534 AW070190 AW893483 AA770330 AA906928 AA906582 AA758746 AA551717 AW063311 AA429538
30	311896	579192_1	AW206447 AI248530 AI084433 AI400976 R16553
	319834	112523_1	AA071267 T65940 T64515 AA071334
	321102	80531_1	AA018306 H38925 AA001221
	321158	410938_1	H79670 H47798 AA700289
	321199	212379_1	N34524 AA305071 AW954803 AA502335 AI433430 AI203597 AW026670 AW265323 AW850787 AA317554 AW993643 AW835572 AW385512 AI334966 W32951 H62656 H53902 R88904 AW835732
35	305528	28832_-3	AA769156
	321270	1662057_1	N59537 N78278 R83560
	314126	177666_1	AA226431 AA226569 AA488748
	320714	743644_1	R91883 AI445591
40	306442	AA976899	
	306446	AA977348	
	306458	AA978186	
	306510	AA988546	
	306557	AA994530	
45	306572	AA995686	
	306582	AA996248	
	306656	AI004024	
	306686	AI015615	
	306751	AI032589	
50	308011	AI439473	
	306892	AI092465	
	308106	AI476803	
	308154	AI500600	
	306956	AI125111	
55	306958	AI125152	
	308213	AI557041	
	308216	AI557135	
	308219	AI557246	
	308588	AI718299	
60	308599	AI719893	
	308643	AI745040	
	308673	AI760864	
	308697	AI767143	
	308778	AI811109	
65	308808	AI818289	
	308875	AI832332	
	308886	AI833240	
	308898	AI858845	
	308966	AI870704	
	308979	AI873111	
70	303011	41689_1	AF090405 AF090407 AF090406
	303077	44060_1	AF163305 AF163307 AF163303
	305016	AA626876	
	305034	AA630128	
75	305072	AA641012	
	305148	AA654070	
	305190	AA665955	
	303978	AW513315	
	303990	AW515465	
80	303998	AW516449	
	303999	AW516611	
	305235	AA670480	
	305312	AA700201	
	305413	AA724659	
	305447	AA737856	
85	321244	29327_1	AF068654 AF068656 AF068655



	305614	AA782866	
	305637	AA806124	
	305639	AA806138	
5	305650	AA807709	
	305690	AA813477	
	305728	AA828209	
	305759	AA835353	
	305792	AA845256	
10	307041	AI144243	
	307091	AI167439	
	307181	AI189251	
	305901	AA872968	
	305910	AA875981	
15	307415	AI242118	
	307426	AI243364	
	307517	AI275055	
	307551	AI281556	
	307561	AI282207	
20	307608	AI290295	
	307691	AI318285	
	307730	AI336092	
	307760	AI342387	
	307764	AI342731	
25	307796	AI350556	
	309045	AI910902	
	309051	AI911975	
	307807	AI351799	
	307808	AI351826	
30	307820	AI355761	
	307852	AI365541	
	309122	AI928178	
	309164	AI937761	
	309177	AI951118	
35	307902	AI380462	
	309299	AW003478	
	309303	AW004823	
	309476	AW129368	
	309532	AW151119	
40	309747	AW264889	
	309769	AW272346	
	309799	AW276964	
	309866	AW299916	
	302679	311853_1	H65022 AA186889
45	309923	AW340684	
	309928	AW341418	
	309931	AW341683	
	309933	AW341936	
	302705	31765_1	U09060 U09061
50	302789	34161_1	AJ245067 AJ245070
	304006	AW517947	
	304024	T03036	
	304026	T03160	
	304028	T03266	
55	304046	T54803	
	304061	T61521	
	304063	T62536	
	302802	34487_1	Y08250 Y08245
	304114	R78946	
60	304155	H68696	
	304203	N56929	
	304234	W81608	
	304348	AA179868	
	304430	AA347682	
65	304456	AA411240	
	304521	AA464716	
	304526	AA476427	
	304607	AA513322	
	304735	AA576453	
70	304760	AA580401	
	306015	AA897116	
	306063	AA906316	
	306065	AA906725	
	306104	AA910956	
75	306109	AA911861	
	306242	AA932805	
	306288	AA936900	
	306396	AA970223	
	330568	NOT_FOUND_entrez	U56244
80	330599	15323_-12	U90437
	331131	genbank_R54797	R54797
	331203	NOT_FOUND_entrez	T82310
	331531	genbank_N51343	N51343
	331547	467396_1	AA828597 N54811
85	332074	genbank_AA599012	AA599012



TABLE 8C shows the genomic position for those Pkeys in Table 8A lacking unigene ID's and accession numbers. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

5	Pkey:	Unique number corresponding to an Eos probeset		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
10	Nt_position:	Indicates nucleotide positions of predicted exons.		
	Pkey	Ref	Strand	Nt_position
	332792	Dunham, I. et.al.	Plus	73381-73768
	332816	Dunham, I. et.al.	Plus	359844-360030
15	332906	Dunham, I. et.al.	Plus	1923101-1923205
	332911	Dunham, I. et.al.	Plus	1961767-1961858
	332912	Dunham, I. et.al.	Plus	1962120-1962246
	332922	Dunham, I. et.al.	Plus	2009620-2009738
	332956	Dunham, I. et.al.	Plus	2510528-2510658
20	332959	Dunham, I. et.al.	Plus	2518145-2518213
	333138	Dunham, I. et.al.	Plus	3369205-3369323
	333139	Dunham, I. et.al.	Plus	3369495-3369571
	333221	Dunham, I. et.al.	Plus	3978070-3978187
25	333380	Dunham, I. et.al.	Plus	4904775-4904846
	333387	Dunham, I. et.al.	Plus	4910935-4910997
	333512	Dunham, I. et.al.	Plus	5560510-5560564
	333524	Dunham, I. et.al.	Plus	5612620-5612780
	333585	Dunham, I. et.al.	Plus	6234778-6234894
	333618	Dunham, I. et.al.	Plus	6562391-6562566
30	333627	Dunham, I. et.al.	Plus	6620584-6620903
	333628	Dunham, I. et.al.	Plus	6629004-6629233
	333650	Dunham, I. et.al.	Plus	6796852-6797128
	333678	Dunham, I. et.al.	Plus	7068223-7068288
	333750	Dunham, I. et.al.	Plus	7608165-7608234
35	333763	Dunham, I. et.al.	Plus	7692491-7692630
	333767	Dunham, I. et.al.	Plus	7694407-7694623
	333768	Dunham, I. et.al.	Plus	7695440-7695697
	333769	Dunham, I. et.al.	Plus	7696625-7696707
	333772	Dunham, I. et.al.	Plus	7706773-7706902
40	333777	Dunham, I. et.al.	Plus	7746805-7746916
	333846	Dunham, I. et.al.	Plus	8008623-8008757
	333884	Dunham, I. et.al.	Plus	8153960-8154161
	333887	Dunham, I. et.al.	Plus	8154882-8155025
45	333891	Dunham, I. et.al.	Plus	8156437-8156709
	333892	Dunham, I. et.al.	Plus	8156825-8157001
	333948	Dunham, I. et.al.	Plus	8583497-8583627
	333954	Dunham, I. et.al.	Plus	6563186-6563335
	333966	Dunham, I. et.al.	Plus	8655643-8655826
	333968	Dunham, I. et.al.	Plus	8681004-8681241
50	334061	Dunham, I. et.al.	Plus	9686941-9687077
	334094	Dunham, I. et.al.	Plus	9889953-9890105
	334113	Dunham, I. et.al.	Plus	10282459-10282597
	334161	Dunham, I. et.al.	Plus	10599033-10599180
	334219	Dunham, I. et.al.	Plus	12716160-12716384
55	334239	Dunham, I. et.al.	Plus	13056569-13056693
	334333	Dunham, I. et.al.	Plus	13603544-13603657
	334378	Dunham, I. et.al.	Plus	13907239-13907370
	334382	Dunham, I. et.al.	Plus	13915866-13916036
	334562	Dunham, I. et.al.	Plus	14987847-14987940
60	334588	Dunham, I. et.al.	Plus	15032740-15032817
	334616	Dunham, I. et.al.	Plus	15176123-15176470
	334633	Dunham, I. et.al.	Plus	15333206-15333305
	334866	Dunham, I. et.al.	Plus	18872214-18872317
	334891	Dunham, I. et.al.	Plus	19299770-19299944
65	334934	Dunham, I. et.al.	Plus	20103970-20104058
	335015	Dunham, I. et.al.	Plus	20682792-20682945
	335120	Dunham, I. et.al.	Plus	21436286-21436384
	335125	Dunham, I. et.al.	Plus	21441390-21441471
	335179	Dunham, I. et.al.	Plus	21634405-21634526
70	335188	Dunham, I. et.al.	Plus	21669118-21669328
	335211	Dunham, I. et.al.	Plus	21774611-21774680
	335361	Dunham, I. et.al.	Plus	22807292-22807445
	335379	Dunham, I. et.al.	Plus	22899306-22899420
	335414	Dunham, I. et.al.	Plus	23235546-23235684
75	335416	Dunham, I. et.al.	Plus	23237354-23237465
	335496	Dunham, I. et.al.	Plus	24164386-24164545
	335497	Dunham, I. et.al.	Plus	24167666-24167869
	335558	Dunham, I. et.al.	Plus	24740167-24740347
	335586	Dunham, I. et.al.	Plus	24990333-24990497
80	335686	Dunham, I. et.al.	Plus	25439839-25439920
	335784	Dunham, I. et.al.	Plus	25942710-25942792
	335823	Dunham, I. et.al.	Plus	26365925-26366004
	335983	Dunham, I. et.al.	Plus	27938968-27939070
	335995	Dunham, I. et.al.	Plus	28009044-28009184
85	336021	Dunham, I. et.al.	Plus	28686482-28686559



	336034	Dunham, I. et.al.	Plus	29014404-29014590
	336038	Dunham, I. et.al.	Plus	29022963-29023165
	336107	Dunham, I. et.al.	Plus	29987731-29987869
5	336632	Dunham, I. et.al.	Plus	983890-985529
	336633	Dunham, I. et.al.	Plus	985591-986221
	336634	Dunham, I. et.al.	Plus	986296-986670
	336635	Dunham, I. et.al.	Plus	987908-988364
	336636	Dunham, I. et.al.	Plus	988418-989185
10	336637	Dunham, I. et.al.	Plus	989276-990813
	336638	Dunham, I. et.al.	Plus	991906-993240
	336659	Dunham, I. et.al.	Plus	1896402-1896478
	336694	Dunham, I. et.al.	Plus	2420546-2420616
	336721	Dunham, I. et.al.	Plus	3371522-3371586
15	336900	Dunham, I. et.al.	Plus	10236423-10236523
	336948	Dunham, I. et.al.	Plus	12692290-12692381
	337028	Dunham, I. et.al.	Plus	16644817-16644942
	337054	Dunham, I. et.al.	Plus	17821742-17821922
	337162	Dunham, I. et.al.	Plus	23478943-23479145
20	337183	Dunham, I. et.al.	Plus	23943606-23943696
	337184	Dunham, I. et.al.	Plus	23973949-23974016
	337268	Dunham, I. et.al.	Plus	28011979-28012034
	337299	Dunham, I. et.al.	Plus	29022656-29022775
	337389	Dunham, I. et.al.	Plus	31401509-31401579
25	337493	Dunham, I. et.al.	Plus	33330760-33330981
	337549	Dunham, I. et.al.	Plus	34474472-34474531
	337755	Dunham, I. et.al.	Plus	3971764-3971900
	337809	Dunham, I. et.al.	Plus	4449069-4449193
	337871	Dunham, I. et.al.	Plus	5443027-5443101
30	337958	Dunham, I. et.al.	Plus	6969162-6969270
	338008	Dunham, I. et.al.	Plus	7697068-7697236
	338033	Dunham, I. et.al.	Plus	8092128-8092271
	338110	Dunham, I. et.al.	Plus	10384481-10384621
	338112	Dunham, I. et.al.	Plus	10391398-10391600
35	338145	Dunham, I. et.al.	Plus	11386629-11386692
	338148	Dunham, I. et.al.	Plus	11448985-11449085
	338179	Dunham, I. et.al.	Plus	12808775-12808833
	338197	Dunham, I. et.al.	Plus	13638107-13638181
	338279	Dunham, I. et.al.	Plus	16168944-16169091
40	338316	Dunham, I. et.al.	Plus	17089711-17089988
	338322	Dunham, I. et.al.	Plus	17132477-17132547
	338357	Dunham, I. et.al.	Plus	18062184-18062402
	338359	Dunham, I. et.al.	Plus	18074402-18074501
	338366	Dunham, I. et.al.	Plus	18252026-18252189
45	338374	Dunham, I. et.al.	Plus	18371200-18371282
	338414	Dunham, I. et.al.	Plus	19345573-19345660
	338418	Dunham, I. et.al.	Plus	19435506-19435596
	338501	Dunham, I. et.al.	Plus	21244713-21244828
	338506	Dunham, I. et.al.	Plus	21221871-21221953
50	338523	Dunham, I. et.al.	Plus	21509763-21509864
	338662	Dunham, I. et.al.	Plus	24404720-24404899
	338804	Dunham, I. et.al.	Plus	27236005-27236108
	338836	Dunham, I. et.al.	Plus	27792166-27792272
	338879	Dunham, I. et.al.	Plus	28410653-28410734
55	338937	Dunham, I. et.al.	Plus	29160655-29160725
	338993	Dunham, I. et.al.	Plus	30077787-30078184
	339047	Dunham, I. et.al.	Plus	30760793-30760968
	339100	Dunham, I. et.al.	Plus	31141580-31141765
	339114	Dunham, I. et.al.	Plus	31456454-31456519
60	339121	Dunham, I. et.al.	Plus	31583467-31583536
	339170	Dunham, I. et.al.	Plus	32216399-32216527
	339293	Dunham, I. et.al.	Plus	33223671-33223819
	332858	Dunham, I. et.al.	Minus	1339607-1339397
	332982	Dunham, I. et.al.	Minus	2628296-2628109
65	332984	Dunham, I. et.al.	Minus	2632606-2632457
	332998	Dunham, I. et.al.	Minus	2711704-2711565
	333058	Dunham, I. et.al.	Minus	3028925-3028811
	333097	Dunham, I. et.al.	Minus	3204124-3204036
	333121	Dunham, I. et.al.	Minus	3308446-3308358
70	333122	Dunham, I. et.al.	Minus	3309596-3309531
	333123	Dunham, I. et.al.	Minus	3310817-3310749
	333140	Dunham, I. et.al.	Minus	3377220-3376309
	333260	Dunham, I. et.al.	Minus	4308400-4308304
	333603	Dunham, I. et.al.	Minus	6466335-6465727
75	333604	Dunham, I. et.al.	Minus	6467090-6466768
	333904	Dunham, I. et.al.	Minus	8217374-8217261
	333906	Dunham, I. et.al.	Minus	8218238-8218063
	334183	Dunham, I. et.al.	Minus	11832582-11832508
	334187	Dunham, I. et.al.	Minus	11921456-11921205
80	334222	Dunham, I. et.al.	Minus	12732417-12732289
	334223	Dunham, I. et.al.	Minus	12734365-12734269
	334255	Dunham, I. et.al.	Minus	13200776-13200692
	334492	Dunham, I. et.al.	Minus	14478333-14478172
	334648	Dunham, I. et.al.	Minus	15363301-15363222
85	334787	Dunham, I. et.al.	Minus	16299093-16298937
	334933	Dunham, I. et.al.	Minus	20078117-20077991



## WO 02/086443

	334945	Dunham, I. et.al.	Minus	20138885-20138637
	334967	Dunham, I. et.al.	Minus	20173311-20173218
	334990	Dunham, I. et.al.	Minus	20341159-20341087
5	335093	Dunham, I. et.al.	Minus	21297367-21297214
	335288	Dunham, I. et.al.	Minus	22304275-22303770
	335289	Dunham, I. et.al.	Minus	22305950-22305708
	335548	Dunham, I. et.al.	Minus	24662773-24662673
	335551	Dunham, I. et.al.	Minus	24679828-24678961
10	335619	Dunham, I. et.al.	Minus	25082677-25082498
	335620	Dunham, I. et.al.	Minus	25092561-25092434
	335621	Dunham, I. et.al.	Minus	25098878-25098767
	335682	Dunham, I. et.al.	Minus	25421215-25421093
	335755	Dunham, I. et.al.	Minus	25763806-25763747
	335814	Dunham, I. et.al.	Minus	26320043-26319845
15	335815	Dunham, I. et.al.	Minus	26320518-26320421
	335835	Dunham, I. et.al.	Minus	26393311-26393245
	335851	Dunham, I. et.al.	Minus	26604863-26604742
	335868	Dunham, I. et.al.	Minus	26711437-26711300
20	335896	Dunham, I. et.al.	Minus	26977639-26977558
	335936	Dunham, I. et.al.	Minus	27360474-27360400
	335948	Dunham, I. et.al.	Minus	27555924-27555788
	336066	Dunham, I. et.al.	Minus	29241080-29240842
	336205	Dunham, I. et.al.	Minus	30477456-30477311
25	336275	Dunham, I. et.al.	Minus	32086675-32086536
	336292	Dunham, I. et.al.	Minus	32818035-32817927
	336331	Dunham, I. et.al.	Minus	33594527-33594371
	336419	Dunham, I. et.al.	Minus	34052568-34052445
	336675	Dunham, I. et.al.	Minus	2020758-2020664
30	336684	Dunham, I. et.al.	Minus	2158060-2157993
	336716	Dunham, I. et.al.	Minus	3259952-3259862
	336798	Dunham, I. et.al.	Minus	5888954-5888757
	337043	Dunham, I. et.al.	Minus	17407330-17407251
	337046	Dunham, I. et.al.	Minus	17610892-17610821
35	337128	Dunham, I. et.al.	Minus	22215251-22215034
	337192	Dunham, I. et.al.	Minus	24591853-24591771
	337194	Dunham, I. et.al.	Minus	24610510-24610359
	337229	Dunham, I. et.al.	Minus	26716579-26716481
	337325	Dunham, I. et.al.	Minus	30015948-30015800
40	337497	Dunham, I. et.al.	Minus	33371317-33371258
	337500	Dunham, I. et.al.	Minus	33376212-33376158
	337603	Dunham, I. et.al.	Minus	1299296-1299194
	337605	Dunham, I. et.al.	Minus	1346555-1346397
	337671	Dunham, I. et.al.	Minus	3260634-3260547
45	337786	Dunham, I. et.al.	Minus	4133203-4133081
	337862	Dunham, I. et.al.	Minus	5347658-5347550
	338083	Dunham, I. et.al.	Minus	9318438-9318301
	338158	Dunham, I. et.al.	Minus	11794465-11794343
	338161	Dunham, I. et.al.	Minus	12124716-12124658
50	338182	Dunham, I. et.al.	Minus	12824919-12824827
	338189	Dunham, I. et.al.	Minus	12878594-12878478
	338199	Dunham, I. et.al.	Minus	13760865-13760780
	338215	Dunham, I. et.al.	Minus	14055447-14055355
	338469	Dunham, I. et.al.	Minus	20520387-20520242
55	338549	Dunham, I. et.al.	Minus	22049171-22049081
	338561	Dunham, I. et.al.	Minus	22311966-22311856
	338671	Dunham, I. et.al.	Minus	24508421-24508346
	338676	Dunham, I. et.al.	Minus	24637427-24637369
	338726	Dunham, I. et.al.	Minus	25926206-25925618
50	338779	Dunham, I. et.al.	Minus	27030151-27029795
	338871	Dunham, I. et.al.	Minus	28301708-28301611
	338872	Dunham, I. et.al.	Minus	28300921-28300790
	338966	Dunham, I. et.al.	Minus	29614876-29614749
	339229	Dunham, I. et.al.	Minus	32722330-32722199
55	339264	Dunham, I. et.al.	Minus	32975145-32975053
	325228	6381940 Plus		2630-2694
	325235	6381943 Minus		162154-162264
	329588	3962484 Plus		1169-1619
	329560	3962491 Plus		2095-2990
70	329541	3983503 Minus		2765-3059
	325328	5866875 Plus		86780-86854
	325340	6017033 Minus		166656-166819
	325373	5866920 Minus		1136686-1136777
	325367	5866920 Minus		922881-922958
75	325389	5866921 Plus		239672-239759
	325436	5866939 Minus		29778-29907
	325498	5866967 Plus		173372-173930
	325471	6017034 Minus		289268-289342
	325557	6056302 Plus		50921-51050
30	325559	6249595 Minus		118590-119172
	325560	6249595 Minus		133794-133981
	325569	6249599 Plus		79927-80217
	325587	6682462 Plus		126724-126967
	325585	6682462 Plus		73476-73574
35	325597	5866992 Plus		1065020-1065089
	325639	5867002 Plus		253525-253608



## WO 02/086443

	325739	5867038	Minus	205138-205269
	325740	5867038	Minus	207533-207690
	325792	6469828	Minus	1018-1176
5	325735	6552447	Minus	269122-269190
	325685	6682468	Plus	117397-117483
	325686	6682468	Plus	118337-118439
	325819	6682490	Minus	130314-130370
	329764	6048195	Minus	109733-109968
10	329703	6065793	Minus	139994-140138
	329643	6448539	Plus	53403-53537
	329816	6624888	Minus	70296-70423
	329860	6687260	Minus	163474-163605
	325883	5867087	Plus	22498-22663
15	325895	5867097	Plus	358317-358476
	325925	5867124	Plus	115749-115962
	325932	5867127	Plus	7369-7441
	325941	5867133	Minus	64228-64402
	325969	5867153	Plus	101911-102081
20	325971	5867153	Plus	105841-106035
	329993	4567166	Minus	101307-101434
	330020	6671887	Plus	172397-172491
	326163	5867168	Minus	7831-8035
	326274	5867171	Minus	410289-410404
25	326025	5867176	Plus	70854-70915
	326046	5867182	Minus	62668-62825
	326099	5867186	Minus	661381-661510
	326108	5867187	Minus	23784-23903
	326165	5867208	Minus	62787-62929
30	326189	5867212	Plus	69288-69413
	326204	5867218	Minus	148088-148200
	326230	5867230	Minus	301868-301972
	330052	4567182	Plus	352560-352963
	330036	6042048	Plus	117120-117216
35	326360	5867293	Plus	13627-13844
	326589	5867320	Plus	22760-22919
	326393	5867341	Plus	41702-41841
	326505	5867435	Minus	8818-8949
	326515	5867439	Plus	36683-36809
40	326592	6138928	Plus	23689-23828
	330107	6015249	Minus	100091-100282
	330106	6015249	Minus	99443-99778
	330100	6015253	Plus	21166-21301
	330093	6015278	Plus	1043-1199
45	330088	6015293	Plus	37517-37638
	330085	6015302	Minus	59613-59770
	330120	6671864	Minus	127553-127656
	330123	6671869	Minus	35311-35406
	326742	5867611	Minus	95187-95248
50	326605	5867637	Plus	24656-24749
	326818	6117831	Minus	15199-15309
	326720	6552456	Plus	84525-84677
	326770	6598307	Minus	513603-513668
	326692	6682502	Plus	117697-117899
55	326693	6682502	Minus	335002-335095
	326983	5867657	Minus	16023-16581
	326991	5867660	Plus	18147-18339
	326936	6004446	Minus	10217-10357
	326964	6469836	Plus	75340-75456
60	327040	6531965	Plus	783670-783817
	327053	6531965	Plus	2247267-2247437
	327075	6531965	Plus	4041318-4041431
	327085	6531965	Plus	4734947-4735069
	327036	6531965	Plus	319951-320040
65	327130	6531976	Plus	20247-22343
	327156	5866841	Minus	2462-2620
	327288	5867481	Plus	48583-48773
	327332	5867516	Minus	56361-56532
	327220	5867525	Minus	65701-65781
70	327224	5867534	Plus	188468-188544
	327321	6249562	Minus	99745-99836
	327361	6552412	Minus	61013-62130
	327396	5867743	Plus	8702-8820
	327414	5867750	Plus	102461-102586
75	327442	5867759	Plus	111483-111618
	327467	5867772	Plus	88030-88151
	327473	5867775	Plus	75101-75181
	327483	5867783	Plus	181573-181662
	327377	5867793	Minus	37610-37676
80	327562	5867804	Minus	343989-344474
	327568	5867811	Minus	46152-46287
	327606	6004463	Plus	200262-200495
	327611	5867868	Minus	175063-175392
	327642	5867891	Minus	2513-2743
85	327654	5867910	Minus	97564-97710
	327734	5867940	Minus	31003-31583



	327775	5867964	Minus	130791-130871
	327796	5867982	Plus	85267-85405
	327840	6249578	Minus	73065-73206
5	330208	6013599	Plus	66517-66931
	330263	6671884	Minus	101503-101634
	328004	5867993	Minus	157407-157887
	328101	5868020	Plus	289920-290014
	328100	5868020	Minus	263545-263635
10	328113	5868024	Minus	80378-80491
	328157	5868064	Plus	73326-73615
	328196	5868080	Minus	16551-16729
	328197	5868081	Minus	42133-42438
	327940	5868197	Minus	95240-95428
15	327984	5868216	Plus	66611-66677
	328021	5902482	Plus	713478-714590
	328068	6117819	Plus	253903-254022
	328264	6381912	Plus	55086-55404
20	330300	2905862	Minus	3246-3302
	328608	5868222	Minus	87770-87953
	328600	5868229	Minus	38889-40010
	328616	5868239	Plus	293920-294224
	328623	5868246	Minus	120020-120126
	328632	5868247	Plus	76734-76853
25	328666	5868254	Minus	778-901
	328698	5868264	Minus	625555-625633
	328700	5868264	Plus	764089-764203
	328708	5868271	Minus	68114-68854
	328735	5868289	Plus	89389-89455
30	328743	5868289	Plus	274638-274726
	328806	5868324	Plus	29408-29684
	328299	5868366	Minus	149708-149889
	328342	5868383	Plus	59955-60094
	328365	5868387	Minus	270724-270798
35	328369	5868388	Plus	75371-75583
	328381	5868392	Plus	662758-662848
	328451	5868425	Minus	217275-217336
	328481	5868449	Minus	8987-9180
	328500	5868464	Plus	59098-59481
40	328530	5868482	Plus	334973-335406
	328664	6004473	Plus	1193739-1193866
	328861	6381928	Minus	108317-108403
	328908	5868493	Plus	117002-117059
	328933	5868500	Plus	771755-771889
45	328934	5868500	Plus	846342-846448
	328949	6456765	Minus	43552-43619
	330313	6042030	Minus	33642-33775
	329005	5868542	Plus	85470-85673
	330366	2944106	Plus	151837-151914
50	330372	6580495	Minus	317461-317688
	329033	5868561	Minus	5390-5479
	329037	5868562	Minus	32466-32562
	329067	5868591	Minus	146417-147652
	329134	5868679	Plus	29959-30018
55	329157	5868687	Minus	145940-146155
	329178	5868704	Plus	179177-179463
	329192	5868716	Plus	166936-167020
	329194	5868716	Minus	304450-304559
	329204	5868720	Minus	3050-3190
60	329224	5868728	Plus	27422-27664
	329228	5868728	Minus	50118-50287
	329288	5868771	Plus	25554-26299
	329337	5868806	Minus	467155-467222
	329011	6682532	Plus	48658-48741



TABLE 9A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer

Table 9A shows about 1312 genes up-regulated in lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) relative to normal body tissues. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 9B show the accession numbers for those Pkey's lacking UnigenelD's for table 9A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 9C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 9A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2
400195			NM_007057*:Homo sapiens ZW10 interactor	1.00	1.00
400205			NM_006265*:Homo sapiens RAD21 (S. pombe)	15.80	396.00
400220			Eos Control	2.28	2.84
400277			Eos Control	7.68	9.72
400285			Eos Control	1.00	1.00
400288	X06256	Hs.149609	integrin, alpha 5 (fibronectin receptor,	1.04	2.24
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.86	74.00
400301	X03635	Hs.1657	estrogen receptor 1	1.00	1.00
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	1.75	1.65
400328	X87344	Hs.180062	transporter 2, ATP-binding cassette, sub	0.87	1.80
400419	AF084545		Target	156.55	253.00
400512			NM_030878*:Homo sapiens cytochrome P450,	1.00	2.00
400517	AF242388		lengsin	3.67	87.00
400560			NM_030878*:Homo sapiens cytochrome P450,	1.00	1.00
400664			NM_002425:Homo sapiens matrix metallopro	20.26	45.00
400665			NM_002425:Homo sapiens matrix metallopro	1.36	1.07
400666			NM_002425:Homo sapiens matrix metallopro	3.26	3.22
400749			NM_003105*:Homo sapiens sortilin-related	1.00	91.00
400763			Target Exon	7.63	24.00
401027			Target Exon	1.00	1.00
401093			C12000586*:gij6330167[dbj]BAA86477.1) (A	1.00	155.00
401203			Target Exon	1.00	86.00
401212			C12000457*:gij7512178[pir]T30337 polypr	1.00	400.00
401411			ENSP00000247172*:HYPOTHETICAL 126.2 kDa	1.00	72.00
401435			C14000397*:gij7499898[pir]T33295 hypoth	1.00	64.00
401464	AF039241		histone deacetylase 5	3.82	49.00
401714			ENSP00000241802*:CDNA FLJ11007 FIS, CLON	2.02	40.00
401747			Homo sapiens keratin 17 (KRT17)	128.43	68.00
401760			Target Exon	1.74	35.00
401780			NM_005557*:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275*:Homo sapiens keratin 15 (KRT1	4.13	2.70
401797			Target Exon	1.44	2.10
401961			NM_021626:Homo sapiens serine carboxypep	1.41	1.86
401985	AF053004		class I cytokine receptor	1.00	177.00
401994			Target Exon	61.84	47.00
402075			ENSP00000251056*:Plasma membrane calcium	1.00	1.00
402260			NM_001436*:Homo sapiens fibrillarin (FBL	1.58	1.39
402265			Target Exon	2.09	35.00
402297			Target Exon	1.00	92.00
402408			NM_030920*:Homo sapiens hypothetical pro	28.87	13.00
402420			C1000823*:gij10432400[emb]CAC10290.1) (A	1.00	1.44
402674			Target Exon	7.44	243.00
402802			NM_001397:Homo sapiens endothelin conver	1.00	70.00
402994			NM_002463*:Homo sapiens myxovirus (influ	1.37	1.43
403137			NM_005381*:Homo sapiens nucleolin (NCL),	1.00	19.00
403306	NM_006825		transmembrane protein (63kD), endoplasmic	1.00	43.00
403329			Target Exon	1.00	61.00
403381			ENSP00000231844*:Ecotropic virus integra	1.00	119.00
403478			NM_022342:Homo sapiens kinesin protein 9	28.13	136.00
403485			C3001813*:gij12737279[ref]XP_012163.1) k	20.23	76.00
403627			Target Exon	6.30	29.33
403715			Target Exon	1.30	35.00
404044			ENSP00000237855*:DJ398G3.2 (NOVEL PROTEI	1.00	54.00
404076			NM_016020*:Homo sapiens CGI-75 protein (	14.29	91.00
404101			C8000950:gij423560[pir]A47318 RNA-bindin	1.00	1.00
404140			NM_006510:Homo sapiens ret finger protei	1.42	1.44
404165			ENSP00000244562:NRH dehydrogenase [quino	1.00	54.00
404185			Target Exon	1.00	117.00
404210			NM_005936:Homo sapiens myeloid/lymphoid	5.93	13.77
404253			NM_021058*:Homo sapiens H2B histone fami	1.00	1.00



	404287		C6001909:gil704441 dbj BAA18909.1  (D298	29.71	42.00
	404298		C6001238*:gil121715 sp P26697 GTA3_CHICK	1.30	1.00
	404347		Target Exon	1.00	1.00
5	404440		NM_021048:Homo sapiens melanoma antigen,	1.00	15.00
	404721		NM_005596*:Homo sapiens nuclear factor I	1.00	60.00
	404794	NM_000078	cholesteryl ester transfer protein, plas	1.07	1.38
	404854		Target Exon	1.61	2.01
	404877		NM_005365:Homo sapiens melanoma antigen,	1.00	1.00
10	404927		Target Exon	1.00	1.00
	404996		Target Exon	1.00	1.00
	405449		CY000047*:gil11427234 ref XP_009399.1  z	1.00	1.00
	405568		NM_031413*:Homo sapiens cat eye syndrome	1.00	78.00
	405572		Target Exon	0.76	1.14
15	405646		C12000200:gil4557225 ref NP_000005.1  al	1.01	1.28
	405676	BE336714	cytochrome c-1	1.13	2.89
	405770		NM_002362:Homo sapiens melanoma antigen,	45.52	37.00
	405932		C15000305:gil3806122 gb AAC69198.1  (AF0	1.99	1.99
	406137		NM_000179*:Homo sapiens mutS (E. coli) h	2.77	2.38
20	406360		Target Exon	1.00	35.00
	406399		NM_003122*:Homo sapiens serine protease	1.00	39.00
	406467		Target Exon	1.00	1.00
	406621	X57809 Hs.181125	immunoglobulin lambda locus	1.41	1.74
	406642	AJ245210	gb:Homo sapiens mRNA for immunoglobulin	2.16	3.91
25	406663	U24683 Hs.293441	immunoglobulin heavy constant mu	2.07	2.93
	406671	AA129547 Hs.285754	met proto-oncogene (hepatocyte growth fa	15.00	51.00
	406673	M34996 Hs.198253	major histocompatibility complex, class	0.98	3.09
	406676	X58399 Hs.81221	Human L2-9 transcript of unrearranged im,	1.30	1.53
	406678	U77534	gb:Human clone 1A11 immunoglobulin varia	1.33	1.45
30	406685	M18728	gb:Human nonspecific crossreacting antig	1.46	2.85
	406687	M31126 Hs.272822	pregnancy specific beta-1-glycoprotein 9	8.61	8.50
	406690	M29540 Hs.220529	carcinoembryonic antigen-related cell ad	226.37	350.00
	406698	X03068 Hs.73931	major histocompatibility complex, class	1.01	2.52
	406815	AA833930 Hs.288036	tRNA isopentenylpyrophosphate transferas	20.25	32.00
35	406851	AA609784	major histocompatibility complex, class	0.75	1.91
	406964	M21305	gb:Human alpha satellite and satellite 3	38.15	1114.00
	406967	M24349	gb:Human parathyroid hormone-like protei	1.00	1.00
	406974	M57293	gb:Human parathyroid hormone-related pep	1.00	1.00
	407103	AA424881 Hs.256301	hypothetical protein MGC13170	1.77	1.10
40	407128	R83312 Hs.237260	EST	1.00	1.00
	407137	T97307	gb:ye53h05.s1 Soares fetal liver spleen	142.70	135.00
	407168	R45175 Hs.117183	ESTs	2.16	18.00
	407239	AA076350 Hs.67846	leukocyte immunoglobulin-like receptor,	1.10	1.57
	407242	M18728	gb:Human nonspecific crossreacting antig	1.12	2.85
45	407244	M10014 Hs.75431	fibrinogen, gamma polypeptide	3.24	15.38
	407289	AA135159 Hs.203349	Homo sapiens cDNA FLJ12149 fis, clone MA	3.53	3.68
	407300	AA102616 Hs.120769	gb:zn43e07.s1 Stratagene HeLa cell s3 93	19.74	73.00
	407366	AF026942 Hs.271530	gb:Homo sapiens cig33 mRNA, partial sequ	0.06	8.25
	407378	AA299264 Hs.57776	ESTs, Moderately similar to I38022 hypot	1.00	26.00
50	407430	AF169351	gb:Homo sapiens protein tyrosine phospho	1.00	25.00
	407453	AJ132087	gb:Homo sapiens mRNA for axonemal dynein	1.00	75.00
	407577	AW131324 Hs.246759	hypothetical protein MGC12538	1.00	1.00
	407634	AW016569 Hs.136414	UDP-GlcNAc:betaGal beta-1,3-N-acetylgluc	111.20	228.00
	407710	AW022727 Hs.23616	ESTs	1.00	28.00
55	407720	AB037776 Hs.38002	KIAA1355 protein	1.89	1.31
	407746	AK001962	hypothetical protein FLJ11100	1.00	1.00
	407756	AA116021 Hs.38260	ubiquitin specific protease 18	4.51	5.00
	407758	D50915 Hs.38365	KIAA0125 gene product	1.00	28.00
	407782	AA608956 Hs.112619	ESTs, Moderately similar to PURKINJE CEL	0.97	1.14
60	407788	BE514982 Hs.38991	S100 calcium-binding protein A2	7.88	3.83
	407790	AI027274 Hs.288941	Homo sapiens cDNA FLJ14866 fis, clone PL	3.63	42.00
	407811	AW190902 Hs.40098	cysteine knot superfamily 1, BMP antagon	89.96	109.00
	407839	AA045144 Hs.161566	ESTs	173.91	108.00
	407944	R34008 Hs.239727	desmocollin 2	111.30	70.00
65	408000	L11690 Hs.620	bullous pemphigoid antigen 1 (230/240kD)	151.17	8.00
	408031	AA081395 Hs.42173	Homo sapiens cDNA FLJ10366 fis, clone NT	9.91	93.00
	408063	BE086548 Hs.42346	calcineurin-binding protein calsarcin-1	195.78	231.00
	408070	AW148852	gb:xf05d05.x1 NCI_CGAP_Brn35 Homo sapien	1.00	1.00
	408101	AW968504 Hs.123073	CDC2-related protein kinase 7	37.84	61.00
70	408122	AI432652 Hs.42824	hypothetical protein FLJ10718	0.85	1.71
	408212	AA297567 Hs.43728	hypothetical protein	5.88	7.91
	408243	Y00787 Hs.624	interleukin 8	4.27	9.98
	408349	BE546947 Hs.44276	homeo box C10	3.79	3.46
	408353	BE439838 Hs.44298	mitochondrial ribosomal protein S17	1.88	1.65
75	408354	AI382803 Hs.159235	ESTs	1.00	73.00
	408369	R38438 Hs.182575	solute carrier family 15 (H??? transport	1.41	16.50
	408380	AF123050 Hs.44532	diubiquitin	15.19	37.22
	408482	NM_000676 Hs.45743	adenosine A2b receptor	1.65	1.19
	408522	AI541214 Hs.46320	Small proline-rich protein SPRK [human,	1.98	1.24
80	408536	AW381532 Hs.135188	ESTs	1.55	1.50
	408545	AW235405 Hs.253690	ESTs	1.00	1.00
	408572	AA055611 Hs.226568	ESTs, Moderately similar to ALU4_HUMAN A	1.00	44.00
	408633	AW963372 Hs.46677	PRO2000 protein	107.16	56.00
	408660	AA525775	ESTs, Moderately similar to PC4259 ferri	1.00	1.00
85	408761	AA057264 Hs.238936	ESTs, Weakly similar to (define not ava	52.24	141.00
	408771	AW732573 Hs.47584	potassium voltage-gated channel, delayed	3.05	109.00



	408783	AF192522	Hs.47701	NPC1 (Niemann-Pick disease, type C1, gen	1.02	1.07
	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	41.19	61.00
	408805	H69912	Hs.48269	vaccinia related kinase 1	24.67	45.00
5	408841	AW438865	Hs.256862	ESTs	1.00	58.00
	408873	AL046017	Hs.182278	calmodulin 2 (phosphorylase kinase, delt	1.00	89.00
	408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
	408992	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.00	1.00
	408996	AI979168	Hs.344096	glycoprotein (transmembrane) nmb	3.71	5.50
10	409015	BE389387	Hs.49767	NM_004553:Homo sapiens NADH dehydrogenas	1.44	1.24
	409038	T97490	Hs.50002	small inducible cytokine subfamily A (Cy	4.28	5.32
	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	112.42	195.00
	409077	AA401369	Hs.190721	ESTs	1.00	17.00
	409093	BE243834	Hs.50441	CGI-04 protein	2.02	1.93
15	409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
	409142	AL136877	Hs.50758	SMC4 (structural maintenance of chromoso	14.87	6.00
	409187	AF154830	Hs.50966	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
	409228	AI654298	Hs.271695	ESTs, Weakly similar to 2109260A B cell	1.22	1.00
	409234	AI879419	Hs.27206	ESTs	1.00	1.00
20	409268	AA625304	Hs.187579	ESTs	11.90	23.00
	409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	168.91	35.00
	409404	BE220053	Hs.129056	ESTs	1.00	1.00
	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	79.74	96.00
25	409430	R21945	Hs.346735	splicing factor, arginine/serine-rich 5	1.45	2.10
	409446	AI561173	Hs.67688	ESTs	1.00	4.00
	409506	NM_006153	Hs.54589	NCK adaptor protein 1	3.97	28.00
	409522	AA075382		gb:zm87b03.s1 Stratagene ovarian cancer	15.98	141.00
	409582	AA401369	Hs.190721	ESTs	1.00	17.00
30	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	292.12	79.00
	409705	M37762	Hs.56023	brain-derived neurotrophic factor	1.00	82.00
	409719	AI769160	Hs.108681	Homo sapiens brain tumor associated prot	1.00	1.00
	409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	0.12	18.12
	409744	AW675258	Hs.56265	Homo sapiens mRNA; cDNA DKFZp586P2321 (f	20.75	51.00
35	409757	NM_001898	Hs.123114	cystatin SN	22.46	15.80
	409866	AW502152		gb:UI-HF-BR0p-ajr-f-11-0-UI.r1 NIH_MGC_5	1.00	1.00
	409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	1.50	1.09
	409902	AI337658	Hs.156351	ESTs	25.92	50.00
	409935	AW511413	Hs.278025	ESTs	2.63	2.11
40	409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	2.17	4.01
	409958	NM_001523	Hs.57697	hyaluronan synthase 1	0.91	2.07
	410001	AB041036	Hs.57771	kallikrein 11	1.04	2.28
	410032	BE065985		gb:RC3-BT0319-120200-014-a09 BT0319 Homo	1.00	58.00
	410037	AB020725	Hs.58009	KIAA0918 protein	1.00	34.00
45	410044	BE566742	Hs.58169	highly expressed in cancer, rich in leuc	1.00	1.00
	410048	W76467	Hs.58218	proline oxidase homolog	1.03	1.44
	410076	T05387	Hs.7991	ESTs	1.12	1.50
	410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
	410153	BE311926	Hs.15830	hypothetical protein FLJ12691	1.00	1.00
50	410166	AK001376	Hs.59346	hypothetical protein FLJ10514	1.00	1.00
	410193	AI132592	Hs.59757	zinc finger protein 281	42.01	51.00
	410274	AA381807	Hs.61762	hypoxia-inducible protein 2	1.72	1.32
	410309	BE043077	Hs.278153	ESTs	1.00	2.00
	410340	AW182833	Hs.112188	hypothetical protein FLJ13149	32.08	75.00
55	410348	AW182663	Hs.95469	ESTs	1.00	1.00
	410407	X66839	Hs.63287	carbonic anhydrase IX	1.40	1.11
	410418	D31382	Hs.63325	transmembrane protease, serine 4	4.30	2.03
	410438	AB037756	Hs.45207	hypothetical protein KIAA1335	1.00	18.00
	410553	AW016824	Hs.255527	hypothetical protein MGC14128	1.34	1.04
60	410555	W27235	Hs.64311	a disintegrin and metalloproteinase doma	23.99	1.41
	410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
	410681	AW246890	Hs.65425	calbindin 1, (28kD)	10.88	18.92
	410781	AI375672	Hs.165028	ESTs	1.00	57.00
	411027	AF072099	Hs.67846	leukocyte immunoglobulin-like receptor,	1.62	3.78
65	411074	X60435	Hs.68137	adenylate cyclase activating polypeptide	1.00	1.15
	411089	AA456454		cell division cycle 2-like 1 (PITSLRE pr	1.56	1.58
	411152	BE069199		gb:QV3-BT0379-010300-105-g03 BT0379 Homo	1.00	84.00
	411248	AA551538	Hs.334605	Homo sapiens cDNA FLJ14408 fis, clone HE	1.82	1.45
	411252	AB018549	Hs.69328	MD-2 protein	7.32	12.74
70	411263	BE297802	Hs.69360	kinesin-like 6 (mitotic centromere-assoc	3.44	2.55
	411365	W76477	Hs.289082	GM2 ganglioside activator protein	1.35	2.02
	411402	BE297855	Hs.69855	NRAS-related gene	1.00	46.00
	411573	AB029000	Hs.70823	KIAA1077 protein	11.40	11.35
	411579	AC005258	Hs.70830	U6 snRNA-associated Sm-like protein LSm7	1.08	1.90
75	411617	AA247994	Hs.90063	neurocalcin delta	1.74	2.57
	411732	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.02	1.00
	411773	NM_006799	Hs.72026	protease, serine, 21 (testisin)	1.34	2.19
	411789	AF245505	Hs.72157	Adlican	2.19	2.79
	411800	N39342	Hs.103042	microtubule-associated protein 1B	23.34	34.00
80	411945	AL033527	Hs.92137	v-myc avian myelocytomatosis viral oncog	1.00	8.00
	412115	AK001763	Hs.73239	hypothetical protein FLJ10901	2.07	1.64
	412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines	118.48	92.00
	412276	BE262621	Hs.73798	macrophage migration inhibitory factor (	1.98	1.49
	412464	T78141	Hs.22826	ESTs, Weakly similar to I55214 salivary	1.16	1.34
85	412530	AA766268	Hs.266273	hypothetical protein FLJ13346	41.52	84.00
	412537	AL031778		nuclear transcription factor Y, alpha	17.90	55.00



	412659	AW753865	Hs.74376	olfactomedin related ER localized protei	14.65	47.00
	412719	AW016610	Hs.816	ESTs	382.46	128.00
	412723	AA648459	Hs.335951	hypothetical protein AF301222	54.90	1.00
5	412811	H06382		ESTs	1.00	11.00
	412817	AL037159	Hs.74619	proteasome (prosome, macropain) 26S subu	1.63	1.42
	412863	AA121673	Hs.59757	zinc finger protein 281	17.63	56.00
	412924	BE018422	Hs.75258	H2A histone family, member Y	1.00	22.00
	413004	T35901	Hs.75117	interleukin enhancer binding factor 2, 4	2.19	2.05
10	413011	AW068115	Hs.821	biglycan	1.22	1.88
	413048	M93221	Hs.75182	mannose receptor, C type 1	0.30	6.23
	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote	3.43	8.71
	413129	AF292100	Hs.104613	RP42 homolog	4.67	4.77
	413142	M81740	Hs.75212	ornithine decarboxylase 1	1.92	2.59
15	413223	AI732182	Hs.191866	ESTs	5.73	27.00
	413248	T64858	Hs.21433	hypothetical protein DKFZp547J036	0.99	1.06
	413273	U75679	Hs.75257	stem-loop (histone) binding protein	1.00	18.00
	413278	BE563085	Hs.833	interferon-stimulated protein, 15 kDa	1.10	1.09
	413281	AA861271	Hs.222024	transcription factor BMAL2	95.94	69.00
20	413364	BE536218	Hs.137516	fidgetin-like 1	1.00	1.00
	413385	M34455	Hs.840	indoleamine-pyrrole 2,3 dioxygenase	0.95	2.09
	413409	AI638418	Hs.1440	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	1.00	1.00
	413453	AA129640	Hs.128065	ESTs	1.00	31.00
	413527	BE250788	Hs.179882	hypothetical protein FLJ12443	1.08	1.46
	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	79.15	114.00
25	413573	AI733859	Hs.149089	ESTs	1.00	1.00
	413582	AW295647	Hs.71331	hypothetical protein MGC5350	8.80	10.00
	413597	AW302885	Hs.117183	ESTs	1.00	1.00
	413690	BE157489		gb:RC1-HT0375-120200-011-e06 HT0375 Homo	1.00	1.00
30	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B	3.16	2.32
	413719	BE439580	Hs.75498	small inducible cytokine subfamily A (Cy	2.88	9.52
	413753	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalinin	144.10	108.00
	413801	M62246	Hs.35406	ESTs, Highly similar to unnamed protein	1.00	17.00
	413833	Z15005	Hs.75573	centromere protein E (312kD)	1.00	1.00
35	413882	AA132973	Hs.184492	ESTs	64.24	148.00
	413926	AA133338	Hs.54310	ESTs	1.00	67.00
	413943	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	43.42	42.00
	413995	BE048146	Hs.75671	syntaxin 1A (brain)	1.23	1.11
	414035	Y00630	Hs.75716	serine (or cysteine) proteinase inhibito	2.02	2.51
40	414142	AW368397	Hs.334485	Homo sapiens cDNA FLJ14438 fis, clone HE	1.00	102.00
	414180	AI863304	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	6.92	77.00
	414245	BE148072	Hs.75850	WAS protein family, member 1	1.00	1.00
	414275	AW970254	Hs.889	Charot-Leyden crystal protein	1.00	59.00
	414317	BE263280	Hs.75888	phosphogluconate dehydrogenase	1.52	1.73
45	414334	AA824298	Hs.21331	hypothetical protein FLJ10036	1.78	1.72
	414341	D80004	Hs.75909	KIAA0182 protein	33.90	151.00
	414368	W70171	Hs.75939	uridine monophosphate kinase	171.60	97.00
	414416	AW409985	Hs.76084	hypothetical protein MGC2721	2.32	1.85
	414430	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1	226.15	66.00
50	414570	Y00285	Hs.76473	insulin-like growth factor 2 receptor	1.64	1.98
	414618	AI204600	Hs.96978	hypothetical protein MGC10764	1.87	72.00
	414675	R79015	Hs.296281	interleukin enhancer binding factor 1	1.51	1.39
	414683	S78296	Hs.76888	hypothetical protein MGC12702	43.61	64.00
	414696	AF002020	Hs.76918	Niemann-Pick disease, type C1	28.63	71.00
55	414711	AI310440	Hs.288735	Homo sapiens cDNA FLJ13522 fis, clone PL	14.86	42.00
	414718	H95348	Hs.107987	ESTs	1.00	5.00
	414732	AW410976	Hs.77152	minichromosome maintenance deficient (S.	1.64	1.44
	414747	U30872	Hs.77204	centromere protein F (350/400kD, mitosi	65.01	74.00
	414761	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2	130.35	121.00
60	414774	X02419	Hs.77274	plasminogen activator, urokinase	2.24	2.19
	414806	D14694	Hs.77329	phosphatidylserine synthase 1	1.63	1.53
	414809	AI434699	Hs.77356	transferrin receptor (p90, CD71)	1.97	2.60
	414812	X72755	Hs.77367	monokine induced by gamma interferon	3.48	10.60
	414825	X06370	Hs.77432	epidermal growth factor receptor (avian	103.22	143.00
65	414839	X63692	Hs.77462	DNA (cytosine-5-)-methyltransferase 1	1.80	1.69
	414883	AA926960		CDC28 protein kinase 1	14.29	10.06
	414907	X90725	Hs.77597	polo (Drosophila)-like kinase	1.95	2.20
	414914	U49844	Hs.77613	ataxia telangiectasia and Rad3 related	3.00	2.90
	414945	BE076358	Hs.77667	lymphocyte antigen 6 complex, locus E	1.02	1.21
70	414972	BE263782	Hs.77695	KIAA0008 gene product	1.00	1.00
	415014	AW954064	Hs.24951	ESTs	1.42	2.84
	415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	34.72	107.00
	415227	AW821113	Hs.72402	ESTs	1.87	49.00
75	415238	R37780	Hs.21422	ESTs	1.00	1.00
	415263	AA948033	Hs.130853	ESTs	1.00	1.00
	415295	R41450	Hs.6546	ESTs	1.00	1.00
	415339	NM_015156	Hs.78398	KIAA0071 protein	51.18	166.00
	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibito	30.84	63.00
80	415674	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,	1.48	1.39
	415709	AA649850	Hs.278558	ESTs	1.00	1.00
	415735	AA704162	Hs.120811	ESTs, Weakly similar to I38022 hypotheti	1.00	72.00
	415799	AA653718	Hs.225841	DKFZP434D193 protein	6.23	31.00
	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
85	415857	AA866115	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	32.51	35.00
	415989	AI267700		ESTs	78.89	1.00



	416018	AW138239	Hs.78977	proprotein convertase subtilisin/kexin t	1.00	1.00
	416065	BE267931	Hs.78996	proliferating cell nuclear antigen	3.35	2.32
	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (	39.03	3.00
5	416177	AA174069	Hs.187607	ESTs	1.00	9.00
	416178	AI808527	Hs.192822	serologically defined breast cancer anti	3.83	3.76
	416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
	416209	AA236776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h	9.70	1.00
	416239	AL038450	Hs.48948	ESTs	83.87	129.00
10	416250	AA581386	Hs.73452	hypothetical protein MGC10791	1.96	2.12
	416322	BE019494	Hs.79217	pyrroline-5-carboxylate reductase 1	2.08	1.73
	416423	H54375	Hs.268921	ESTs	1.00	89.00
	416448	L13210	Hs.79339	lectin, galactoside-binding, soluble, 3	1.28	1.54
	416498	U33632	Hs.79351	potassium channel, subfamily K, member 1	27.29	67.00
15	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
	416661	AA634543	Hs.79440	IGF-II mRNA-binding protein 3	9.96	5.00
	416722	AA354604	Hs.122546	hypothetical protein FLJ23017	3.68	33.00
	416819	U77735	Hs.80205	pim-2 oncogene	1.59	1.84
	416936	N21352	Hs.42987	ESTs, Weakly similar to S21348 probable	1.00	1.00
20	417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
	417061	AI675944	Hs.188691	Homo sapiens cDNA FLJ12033 fis, clone HE	32.95	156.00
	417079	U65590	Hs.81134	interleukin 1 receptor antagonist	3.91	4.93
	417218	AA129547	Hs.285754	met proto-oncogene (hepatocyte growth fa	1.00	51.00
	417233	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	3.38	2.05
25	417308	H60720	Hs.81892	KIAA0101 gene product	82.94	25.36
	417315	AI080042	Hs.180450	ribosomal protein S24	106.61	121.00
	417324	AW265494		ESTs	1.20	1.28
	417366	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)	8.97	3.27
	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	2.59	1.82
30	417428	N87579	Hs.278871	gb:LL2030F Human fetal heart, Lambda ZAP	1.00	52.00
	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	304.75	173.00
	417466	AI681547	Hs.59457	hypothetical protein FLJ22127	1.24	1.34
	417512	AI979168	Hs.344096	glycoprotein (transmembrane) nmb	2.14	5.50
	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated	2.66	1.68
35	417542	J04129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
	417576	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	42.76	51.00
	417715	AW969587	Hs.86366	ESTs	6.35	2.75
	417720	AA205625	Hs.208067	ESTs	113.31	56.00
	417791	AW965339	Hs.111471	ESTs	39.98	16.00
40	417830	AW504786	Hs.122579	hypothetical protein FLJ10461	2.61	31.00
	417866	AW067903	Hs.82772	collagen, type XI, alpha 1	2.35	2.44
	417900	BE250127	Hs.82906	CDC20 (cell division cycle 20, S. cerevi	1.52	1.11
	417933	X02308	Hs.82962	thymidylate synthetase	4.74	2.55
	417944	AU077196	Hs.82985	collagen, type V, alpha 2	3.61	5.21
45	417975	AA641836	Hs.30085	hypothetical protein FLJ23186	12.49	38.00
	417991	AA731452	Hs.190008	ESTs	1.00	26.00
	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	3.02	2.12
	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	187.59	1.00
	418054	NM_002318	Hs.83354	lysyl oxidase-like 2	2.85	2.63
50	418057	NM_012151	Hs.83363	coagulation factor VIII-associated (intr	1.54	1.69
	418113	AI272141	Hs.83484	SRY (sex determining region Y)-box 4	6.82	5.22
	418140	BE613836	Hs.83551	microfibrillar-associated protein 2	1.26	1.46
	418203	X54942	Hs.83758	CDC28 protein kinase 2	134.19	144.00
	418207	C14685	Hs.34772	ESTs	1.00	1.00
55	418216	AA662240	Hs.283099	AF15q14 protein	64.66	61.00
	418236	AW994005	Hs.337534	ESTs	18.53	147.00
	418249	H89226	Hs.34892	KIAA1323 protein	30.53	106.00
	418281	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9	1.00	3.00
	418283	S79895	Hs.83942	calthepsin K (pseudosynostosis)	3.96	5.16
60	418300	AI433074	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	3.18	2.91
	418322	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	11.96	6.68
	418327	U70370	Hs.84136	paired-like homeodomain transcription fa	9.23	2.22
	418345	AJ001696	Hs.241407	serine (or cysteine) proteinase inhibito	1.00	1.00
	418379	AA218940	Hs.137516	fidgetin-like 1	21.68	44.00
65	418397	NM_001269	Hs.84746	chromosome condensation 1	1.00	8.00
	418403	D86978	Hs.84790	KIAA0225 protein	16.91	18.98
	418462	BE001596	Hs.85266	integrin, beta 4	1.56	1.16
	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	3.22	2.38
	418506	AA084248	Hs.85339	G protein-coupled receptor 39	2.66	2.22
70	418526	BE019020	Hs.85838	solute carrier family 16 (monocarboxylic	2.04	2.21
	418538	BE244323	Hs.85951	exportin, tRNA (nuclear export receptor	1.33	37.00
	418543	NM_005329	Hs.85962	hyaluronan synthase 3	1.04	1.23
	418574	N28754		M-phase phosphoprotein 9	48.60	85.00
	418592	X99226	Hs.284153	Fanconi anemia, complementation group A	18.24	26.00
75	418641	BE243136	Hs.86947	a disintegrin and metalloproteinase doma	1.19	1.41
	418661	NM_001949	Hs.1189	E2F transcription factor 3	29.05	43.00
	418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
	418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
	418686	Z36830	Hs.87268	annexin A8	1.54	1.98
80	418689	AI360883	Hs.274448	hypothetical protein FLJ11029	1.19	1.04
	418712	Z42183		gb:HSC0BF041 normalized infant brain cDN	1.00	12.00
	418727	AA227609	Hs.94834	ESTs	1.00	49.00
	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino	49.85	1.00
	418819	AA228776	Hs.191721	ESTs	1.00	140.00
85	418830	BE513731	Hs.88959	hypothetical protein MGC4816	20.97	23.00
	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	57.09	35.00



	418971	AA360392	Hs.87113	ESTs	1.00	12.00
	418973	AA233056	Hs.191518	ESTs	4.89	28.00
	419078	M93119	Hs.89584	insulinoma-associated 1	1.00	10.00
5	419079	AW014836	Hs.18844	ESTs	1.09	1.98
	419080	AW150835	Hs.18878	hypothetical protein FLJ21620	2.06	1.68
	419088	AI538323	Hs.52620	integrin, beta 8	15.60	51.00
	419092	J05581	Hs.89603	mucin 1, transmembrane	1.11	1.83
	419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
10	419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor t	1.10	1.14
	419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy	3.18	2.43
	419288	AA256106	Hs.87507	ESTs	1.00	34.00
	419335	AW960146	Hs.284137	hypothetical protein FLJ12888	1.00	8.00
15	419354	M62839	Hs.1252	apolipoprotein H (beta-2-glycoprotein I)	22.63	54.00
	419359	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)	2.50	1.98
	419423	D26488	Hs.90315	KIAA0007 protein	1.00	7.00
	419443	D62703		gb:HUM316G10B Clontech human aorta polyA	1.00	12.00
	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	1.64	1.84
20	419474	AW968619	Hs.155849	ESTs	13.63	62.00
	419485	AA489023	Hs.99807	ESTs, Weakly similar to unnamed protein	4.27	2.26
	419488	AA316241	Hs.90691	nucleophosmin/nucleoplasmin 3	3.66	3.63
	419502	AU076704		fibrinogen, A alpha polypeptide	13.05	115.00
	419539	AF070590	Hs.90869	Homo sapiens clones 24622 and 24623 mRNA	74.60	117.00
25	419556	U29615	Hs.91093	chitinase 1 (chitotriosidase)	1.47	4.98
	419569	AI971651	Hs.91143	jagged 1 (Alagille syndrome)	1.00	4.00
	419594	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	94.30	94.00
	419703	AI793257	Hs.128151	ESTs	15.26	50.00
	419721	NM_001650	Hs.288650	aquaporin 4	1.00	191.00
30	419729	AA586442	Hs.21411	gb:no53a03.s1 NCI_CGAP_SS1 Homo sapiens	1.00	59.00
	419741	NM_007019	Hs.93002	ubiquitin carrier protein E2-C	2.02	1.08
	419745	AF042001	Hs.93005	slug (chicken homolog), zinc finger prot	1.00	1.00
	419752	AA249573	Hs.152618	ESTs, Moderately similar to ZN91_HUMAN Z	29.87	77.00
	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	50.99	214.00
35	419936	AI792788		gb:ol91d05.y5 NCI_CGAP_Kid5 Homo sapiens	1.00	1.00
	419937	AB040959	Hs.93836	DKFZP434N014 protein	1.64	2.47
	419983	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f	15.72	94.00
	420005	AW271106	Hs.133294	ESTs	3.15	1.43
	420047	AI478658	Hs.94631	brefeldin A-inhibited guanine nucleotide	12.45	39.00
40	420058	AK001423	Hs.94694	Homo sapiens cDNA FLJ10561 fis, clone NT	1.00	117.00
	420162	BE378432	Hs.95577	cyclin-dependent kinase 4	1.43	1.21
	420251	AW374968	Hs.348112	Human DNA sequence from clone RP5-1103G7	2.35	3.23
	420259	AF004884	Hs.96253	calcium channel, voltage-dependent, P/Q	0.77	1.15
	420281	AI623693	Hs.323494	ESTs	45.04	54.00
45	420309	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S	49.22	31.00
	420332	NM_001756	Hs.1305	serine (or cysteine) proteinase inhibito	0.05	2.82
	420380	AA640891	Hs.102406	ESTs	0.99	2.74
	420462	AF050147	Hs.97932	chondromodulin I precursor	1.00	1.00
	420520	AK001978	Hs.98510	similar to rab11-binding protein	49.74	133.00
50	420552	AK000492	Hs.98806	hypothetical protein	94.65	88.00
	420560	AW207748	Hs.59115	ESTs	1.00	17.00
	420610	AI683183	Hs.99348	distal-less homeo box 5	1.00	13.00
	420689	H79979	Hs.88678	ESTs	50.09	95.00
	420721	AA927802	Hs.159471	ZAP3 protein	1.00	31.00
55	420759	T11832	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	1.00	48.00
	420783	AI659838	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
	420900	AL045633	Hs.44269	ESTs	2.24	7.00
	420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
	421002	AF116030	Hs.100932	transcription factor 17	1.00	27.00
60	421027	AA761198	Hs.55254	ESTs	2.87	38.00
	421037	AI684808	Hs.197653	ESTs	1.00	46.00
	421041	N36914	Hs.14691	ESTs, Moderately similar to I38022 hypot	1.00	98.00
	421073	NM_004689	Hs.101448	metastasis associated 1	1.34	1.46
	421110	AJ250717	Hs.1355	cathepsin E	119.47	427.00
65	421133	AA401369	Hs.190721	ESTs	1.10	17.00
	421150	AI913562	Hs.189902	ESTs	1.45	1.63
	421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
	421307	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f	1.37	1.10
	421316	AA287203	Hs.324728	SMA5	1.00	21.00
70	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	1.92	3.94
	421451	AA291377	Hs.50831	ESTs	5.89	14.00
	421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
	421506	BE302796	Hs.105097	thymidine kinase 1, soluble	1.56	1.08
	421508	NM_004833	Hs.105115	absent in melanoma 2	5.11	5.23
75	421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase 1, I	1.00	3.00
	421524	AA312082	Hs.105445	GDNF family receptor alpha 1	2.63	10.58
	421526	AL080121	Hs.105460	DKFZP564O0823 protein	1.46	1.88
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	30.21	50.32
	421574	AJ000152	Hs.105924	defensin, beta 2	1.67	1.74
80	421582	AI910275		trefoil factor 1 (breast cancer, estroge	1.23	1.00
	421633	AF121860	Hs.106260	sorting nexin 10	1.00	116.00
	421659	NM_014459	Hs.106511	protocadherin 17	0.05	6.33
	421677	H64092	Hs.38282	ESTs	1.31	1.42
	421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	1.41	1.20
85	421773	W69233	Hs.112457	ESTs	1.12	1.14
	421777	BE562088	Hs.108196	HSPC037 protein	1.97	1.29



	421800	AA298151	Hs.222969	ESTs	1.03	1.30
	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.88	1.59
	421896	N62293	Hs.45107	ESTs	11.84	22.80
5	421928	AF013758	Hs.109643	polyadenylate binding protein-interactin	45.89	90.00
	421931	NM_000814	Hs.1440	gamma-aminobutyric acid (GABA) A recepto	1.13	1.49
	421948	L42583	Hs.334309	keratin 6A	51.83	20.25
	421975	AW961017	Hs.6459	hypothetical protein FLJ11856	1.17	1.15
	422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
10	422094	AF129535	Hs.272027	F-box only protein 5	67.61	62.00
	422095	AI868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
	422109	S73265	Hs.1473	gastrin-releasing peptide	4.18	95.50
	422128	AW881145		gb:QV0-OT0033-010400-182-a07 OT0033 Homo	40.89	71.00
	422129	AU076635	Hs.1478	serine (or cysteine) proteinase inhibito	1.13	1.38
15	422134	AW179019	Hs.112110	mitochondrial ribosomal protein L42	41.59	96.00
	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias	3.29	1.68
	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6	4.93	5.73
	422282	AF019225	Hs.114309	apolipoprotein L	1.49	1.71
20	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	25.99	10.91
	422310	AA316622	Hs.98370	cytochrome P450, subfamily IIS, polypept	1.54	1.41
	422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
	422330	D30783	Hs.115263	epiregulin	1.00	112.00
	422364	AF067800	Hs.115515	C-type (calcium dependent, carbohydrate-	9.39	60.00
25	422406	AF025441	Hs.116206	Opa-interacting protein 5	18.33	53.00
	422424	AI186431	Hs.296638	prostate differentiation factor	1.71	3.21
	422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
	422487	AJ010901	Hs.198267	mucin 4, tracheobronchial	73.68	35.54
	422511	AU076442	Hs.117938	collagen, type XVII, alpha 1	173.97	26.00
30	422515	AW500470	Hs.117950	multifunctional polypeptide similar to S	4.68	2.92
	422656	AI870435	Hs.1569	LIM homeobox protein 2	1.00	1.00
	422737	M26939	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	3.89	4.55
	422756	AA441787	Hs.119689	glycoprotein hormones, alpha polypeptide	1.05	1.46
	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	3.88	1.53
35	422809	AK001379	Hs.121028	hypothetical protein FLJ10549	99.56	53.00
	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
	422938	NM_001809	Hs.1594	centromere protein A (17kD)	70.46	61.00
	422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	77.74	3.00
	422960	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	5.88	8.55
40	422963	AA401369	Hs.190721	ESTs	171.41	17.00
	422976	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e	2.12	1.62
	422981	AF026445	Hs.122752	TATA box binding protein (TBP)-associate	10.49	35.00
	422986	AA319777	Hs.221974	ESTs	12.40	32.47
	423034	AL119930		gb:DKFZp761A092_r1 761 (synonym: hamy2)	16.41	60.00
45	423049	X59373	Hs.188023	ESTs, Moderately similar to HXDA_HUMAN H	1.00	1.00
	423081	AF262992	Hs.123159	sperm associated antigen 4	1.82	2.96
	423184	NM_004428	Hs.1624	ephrin-A1	1.14	1.53
	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys	2.14	1.69
	423248	AA380177	Hs.125845	ribulose-5-phosphate-3-epimerase	7.18	14.00
50	423309	BE006775	Hs.126782	sushi-repeat protein	21.90	64.00
	423361	AW170055	Hs.47628	ESTs	1.00	1.00
	423453	AW450737	Hs.128791	CGI-09 protein	55.52	66.00
	423511	AF036329	Hs.129715	gonadotropin-releasing hormone 2	0.88	1.17
	423516	AB007933	Hs.129729	ligand of neuronal nitric oxide synthase	1.76	5.40
55	423551	AA327598	Hs.233785	ESTs	3.54	4.33
	423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
	423575	C18863	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	38.88	70.00
	423624	AI807408	Hs.166368	ESTs	1.00	67.00
	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
60	423642	AW452650	Hs.157148	hypothetical protein MGC13204	19.14	58.00
	423662	AA642452	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	3.61	13.57
	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	240.73	40.00
	423698	AA329796	Hs.1098	DKFZp434J1813 protein	1.00	59.00
	423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
65	423761	NM_006194	Hs.132576	paired box gene 9	1.00	1.00
	423787	AJ295745	Hs.236204	nuclear pore complex protein	7.18	6.64
	423816	AF151064		hypothetical protein	1.00	44.00
	423826	U20325	Hs.1707	cocaine- and amphetamine-regulated trans	1.00	1.00
	423849	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	1.00	1.00
70	423887	AL080207	Hs.134585	DKFZP434G232 protein	1.00	1.00
	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f	31.33	31.00
	423954	AW753164	Hs.288604	KIAA1632 protein	5.81	10.87
	423961	D13666	Hs.136348	osteoblast specific factor 2 (fasciclin	3.55	3.30
	424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	233.42	68.00
75	424016	AW163729	Hs.6140	hypothetical protein MGC15730	0.93	1.01
	424028	AF055084	Hs.153692	Homo sapiens cDNA FLJ14354 fis, clone Y7	21.30	52.00
	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito	1.00	1.00
	424086	AI351010	Hs.102267	lysyl oxidase	21.91	70.00
	424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
80	424120	T80579	Hs.290270	ESTs	1.00	1.00
	424165	AW582904	Hs.142255	islet amyloid polypeptide	1.00	34.00
	424200	AA337221		gb:EST41944 Endometrial tumor Homo sapie	13.06	48.00
	424279	L29306	Hs.171814	tryptophan hydroxylase (tryptophan 5-mon	1.00	1.00
	424308	AW975531	Hs.154443	minichromosome maintenance deficient (S.	164.58	87.00
85	424326	NM_014479	Hs.145296	disintegrin protease	53.72	302.00
	424340	AA339036	Hs.7033	ESTs	0.88	1.15



	424351	BE622117	Hs.145567	hypothetical protein	0.93	1.03
	424364	AW383226	Hs.201189	ESTs, Weakly similar to G01763 atrophin-	7.02	3.24
	424381	AA285249	Hs.146329	protein kinase Chk2	95.55	92.00
5	424411	NM_005209	Hs.146549	crystallin, beta A2	1.63	3.25
	424420	BE614743	Hs.146688	prostaglandin E synthase	1.63	1.33
	424441	X14850	Hs.147097	H2A histone family, member X	1.82	1.29
	424502	AF242388	Hs.149585	lengsin	1.00	1.00
	424503	X06256	Hs.149609	integrin, alpha 5 (fibronectin receptor,	1.02	2.24
10	424513	BE385864	Hs.149894	mitochondrial translational initiation f	1.00	17.00
	424539	L02911	Hs.150402	Activin A receptor, type I (ACVR1) (ALK	32.46	108.00
	424568	AF005418	Hs.150595	cytochrome P450, subfamily XXVIA, polype	3.40	2.58
	424602	AK002055	Hs.151046	hypothetical protein FLJ11193	31.87	25.00
	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub	3.58	2.37
15	424645	NM_014682	Hs.151449	KIAA0535 gene product	1.00	1.00
	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	2.12	2.23
	424717	AW992292	Hs.152213	wingless-type MMTV integration site fami	1.00	1.00
	424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
	424840	D79987	Hs.153479	extra spindle poles, S. cerevisiae, homo	2.65	1.30
20	424867	A1024860	Hs.153591	Not56 (D. melanogaster)-like protein	1.23	1.05
	424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
	424979	D87989	Hs.154073	UDP-galactose transporter related	1.36	1.35
	424999	AW953120		gb:EST365190 MAGE resequences, MAGB Homo	1.24	1.41
	425048	H05468	Hs.164502	ESTs	1.00	11.00
25	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	7.46	87.00
	425081	X74794	Hs.154443	minichromosome maintenance deficient (S.	2.52	3.82
	425118	AU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase	4.84	4.03
	425159	NM_004341	Hs.154868	carbamoyl-phosphate synthetase 2, aspart	3.62	2.73
	425202	AW962282	Hs.152049	ESTs, Weakly similar to I38022 hypotheli	1.00	53.00
30	425234	AW152225	Hs.165909	ESTs, Weakly similar to I38022 hypotheli	100.77	44.00
	425236	AW067800	Hs.155223	stanniocalcin 2	3.30	2.90
	425245	A1751768	Hs.155314	KIAA0095 gene product	1.91	2.32
	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	1.41	1.49
	425266	J00077	Hs.155421	alpha-fetoprotein	1.00	68.00
35	425274	BE281191	Hs.155462	minichromosome maintenance deficient (mi	1.97	1.63
	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	141.49	123.00
	425349	AA425234	Hs.79886	ribose 5-phosphate isomerase A (ribose 5	1.00	84.00
	425371	D49441	Hs.155981	mesothelin	0.87	1.59
	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	14.90	5.76
40	425420	BE536911	Hs.234545	hypothetical protein NUF2R	1.00	1.00
	425424	NM_004954	Hs.157199	ELKL motif kinase	10.58	9.74
	425483	AF231022	Hs.158159	FAT tumor suppressor (Drosophila) homolo	1.74	1.40
	425566	AW162943	Hs.250618	UL16 binding protein 2	1.49	1.14
	425580	L11144	Hs.1907	galanin	53.29	233.00
45	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
	425692	D90041	Hs.155956	N-acetyltransferase 1 (arylamine N-acety	1.00	55.00
	425695	NM_005401	Hs.159238	protein tyrosine phosphatase, non-recept	1.00	10.00
	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.00	41.00
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	1.00	48.00
50	425810	A1923627	Hs.31903	ESTs	27.39	98.00
	425811	AL039104	Hs.159557	karyopherin alpha 2 (RAG cohort 1, impor	1.99	1.58
	425849	AI077288	Hs.296323	serum/glucocorticoid regulated kinase	71.16	3.42
	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	1.35	1.34
	426067	AA401369	Hs.190721	ESTs	1.01	17.00
55	426088	AF038007	Hs.166196	ATPase, Class I, type 8B, member 1	26.26	47.00
	426215	AW067800	Hs.155223	stanniocalcin 2	1.91	2.90
	426227	U67058	Hs.154299	Human proteinase activated receptor-2 mR	22.40	25.00
	426269	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f	1.00	1.00
	426283	NM_003937	Hs.169139	kynureninase (L-kynurenine hydrolase)	91.39	229.00
60	426329	AL389951	Hs.271623	nucleoporin 50kD	4.34	4.08
	426427	M86699	Hs.169840	TTK protein kinase	7.02	1.00
	426432	AF001601	Hs.169857	paraoxonase 2	1.16	1.68
	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu	2.59	1.71
	426459	AF151812	Hs.169992	hypothetical 43.2 Kd protein	1.56	1.66
65	426471	M22440	Hs.170009	transforming growth factor, alpha	20.60	26.00
	426496	D31765	Hs.170114	KIAA0061 protein	9.81	22.00
	426501	AA401369	Hs.190721	ESTs	19.23	17.00
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	103.74	41.00
	426536	A1949749	Hs.44441	ESTs	4.65	23.00
70	426572	AB037783	Hs.170623	hypothetical protein FLJ11183	1.00	43.00
	426682	AV660038	Hs.2056	UDP glycosyltransferase 1 family, polype	160.06	8.00
	426691	NM_006201	Hs.171834	PCTAIRE protein kinase 1	1.51	1.35
	426746	J03626	Hs.2057	uridine monophosphate synthetase (orotat	2.13	1.68
	426752	X69490	Hs.172004	titin	0.02	5.14
75	426784	U03749	Hs.172216	chromogranin A (parathyroid secretory pr	1.72	1.71
	426807	AA385315	Hs.156682	ESTs	1.30	1.64
	426812	AF105365	Hs.172613	solute carrier family 12 (potassium/chlo	1.47	1.53
	426814	AF036943	Hs.172619	myelin transcription factor 1-like	1.00	1.00
	426831	BE296216	Hs.172673	S-adenosylhomocysteine hydrolase	1.51	1.25
80	426897	AA401369	Hs.190721	ESTs	141.56	17.00
	426925	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	32.61	38.00
	426935	NM_000088	Hs.172928	collagen, type I, alpha 1	2.65	3.16
	426964	AA393739	Hs.287416	Homo sapiens cDNA FLJ11439 fis, clone HE	1.97	3.49
	426966	A1493134		sclerostin	1.00	1.00
85	426991	AK001536		Homo sapiens cDNA FLJ10674 fis, clone NT	3.39	2.28
	427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00



	427239	BE270447	Hs.174070	ubiquitin carrier protein	1.58	1.05
	427260	AA663848		gb:ae70b06.s1 Stratagene schizo brain S1	1.34	1.60
	427281	AA906147	Hs.102869	ESTs	1.00	66.00
5	427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
	427354	T57896	Hs.191095	ESTs	1.17	1.95
	427356	AW023482	Hs.97849	ESTs	7.31	41.00
	427376	AA401533	Hs.19440	ESTs	1.00	57.00
	427383	NM_005411	Hs.177582	surfactant, pulmonary-associated protein	0.42	1.32
10	427427	AF077345	Hs.177936	lectin, superfamily member 1 (cartilage-	1.00	20.00
	427441	AA412605	Hs.343879	SPANX family, member C	1.00	1.00
	427445	X80818	Hs.178078	glutamate receptor, metabotropic 4	0.97	1.03
	427505	AA361562	Hs.178761	26S proteasome-associated pad1 homolog	4.60	4.04
	427510	Z47542	Hs.179312	small nuclear RNA activating complex, po	22.00	45.00
	427528	AU077143	Hs.179565	minichromosome maintenance deficient (S.	97.45	92.00
15	427546	AA188763	Hs.36793	hypothetical protein FLJ23188	1.50	3.24
	427562	R56424	Hs.26534	ESTs	6.81	40.00
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	69.91	62.00
	427660	AI741320	Hs.114121	Homo sapiens cDNA: FLJ23228 fis, clone C	2.70	49.00
	427666	AI791495	Hs.180142	calmodulin-like skin protein	1.37	1.88
20	427668	AA298760	Hs.180191	hypothetical protein FLJ14904	29.55	67.00
	427677	NM_007045	Hs.180296	FGFR1 oncogene partner	3.52	2.63
	427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
	427711	M31659	Hs.180408	solute carrier family 25 (mitochondrial	15.84	70.00
	427719	AI393122	Hs.134726	ESTs	7.03	4.52
25	427722	AK000123	Hs.180479	hypothetical protein FLJ20116	2.92	1.74
	427747	AW411425	Hs.180655	serine/threonine kinase 12	1.76	1.26
	427912	AL022310	Hs.181097	tumor necrosis factor (ligand) superfami	9.63	59.00
	427961	AW293165	Hs.143134	ESTs	41.97	118.00
30	428004	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	23.82	1.00
	428023	AL038843		Homo sapiens cDNA: FLJ23602 fis, clone L	1.40	1.33
	428046	AW812795	Hs.337534	ESTs, Moderately similar to I38022 hypot	96.28	167.00
	428093	AW594506	Hs.104830	ESTs	1.25	1.29
	428098	AU077258	Hs.182429	protein disulfide isomerase-related prot	1.86	1.60
	428129	AI244311	Hs.26912	ESTs	1.00	42.00
35	428169	AI928984	Hs.182793	golgi phosphoprotein 2	2.76	2.11
	428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	85.59	181.00
	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	8.57	21.64
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	7.77	15.90
40	428434	AI909935	Hs.65551	Homo sapiens, Similar to DNA segment, Ch	0.58	1.43
	428450	NM_014791	Hs.184339	KIAA0175 gene product	237.53	204.00
	428471	X57348	Hs.184510	stratifin	6.00	4.60
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	56.54	16.00
	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	3.53	2.15
45	428505	AL035461	Hs.2281	chromogranin B (secretogranin 1)	1.00	1.00
	428532	AF157326	Hs.184786	TBP-interacting protein	1.00	58.00
	428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	1.00	1.00
	428698	AA852773	Hs.334838	KIAA1866 protein	187.37	255.00
50	428728	NM_016625	Hs.191381	hypothetical protein	47.24	80.00
	428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
	428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
	428771	AB028992	Hs.193143	KIAA1069 protein	1.98	92.00
	428801	AW277121	Hs.254881	ESTs	1.67	6.15
55	428810	AF068236	Hs.193788	nitric oxide synthase 2A (inducible, hep	1.03	1.27
	428839	AI767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	124.17	43.00
	428845	AL157579	Hs.153610	KIAA0751 gene product	1.00	1.00
	428959	AF100779	Hs.194680	WNT1 inducible signaling pathway protein	15.16	27.00
	428969	AF120274	Hs.194689	artemin	1.36	1.24
60	429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like	0.97	3.31
	429065	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	6.82	16.47
	429164	AI688663	Hs.116586	ESTs	19.08	67.00
	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
	429183	AB014604	Hs.197955	KIAA0704 protein	79.72	104.00
65	429201	X03178	Hs.198246	group-specific component (vitamin D bind	1.00	1.00
	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	1.33	1.09
	429220	AW207206		ESTs	1.00	7.00
	429228	AI553633	Hs.326447	ESTs	39.47	29.25
	429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
70	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
	429276	AF056085	Hs.198612	G protein-coupled receptor 51	3.70	142.00
	429359	W00482	Hs.2399	matrix metalloproteinase 14 (membrane-in	1.30	1.94
	429412	NM_006235	Hs.2407	POU domain, class 2, associating factor	94.09	86.00
	429413	NM_014058	Hs.201877	DESC1 protein	41.91	10.00
75	429486	AF155827	Hs.203963	hypothetical protein FLJ10339	12.19	1.00
	429504	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	1.61	1.08
	429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
	429547	AA401369	Hs.190721	ESTs	1.06	17.00
	429551	AW450624	Hs.220931	ESTs	2.89	65.00
80	429563	BE619413	Hs.2437	eukaryotic translation initiation factor	1.49	1.37
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	61.86	100.00
	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	1.59	1.69
	429612	AF062649	Hs.252587	pituitary tumor-transforming 1	2.78	1.74
	429616	AI982722	Hs.120845	ESTs	1.00	1.00
85	429656	X05608	Hs.211584	neurofilament, light polypeptide (68kD)	1.00	4.00



5	429663	M68874	Hs.211587	phospholipase A2, group IVA (cytosolic,	69.95	104.00
	429736	AF125304	Hs.212680	tumor necrosis factor receptor superfam	1.25	1.21
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	1.00	7.00
	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	11.80	1.00
	429918	AW873986	Hs.119383	ESTs	1.00	78.00
10	429978	AA249027		ribosomal protein S6	1.98	3.09
	429986	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	1.00	48.00
	430044	AA464510	Hs.152812	ESTs	69.27	59.00
	430114	AA847744	Hs.99640	ESTs	1.00	1.00
	430134	BE380149	Hs.105223	ESTs, Weakly similar to T33188 hypotheti	1.00	51.00
15	430147	R60704	Hs.234434	hairy/enhancer-of-split related with YRP	1.10	2.22
	430287	AW182459	Hs.125759	ESTs, Weakly similar to LEU5_HUMAN LEUKE	1.00	127.00
	430294	AI538226	Hs.32976	guanine nucleotide binding protein 4	3.80	1.47
	430300	U60805	Hs.238648	oncostatin M receptor	1.00	35.00
	430315	NM_004293	Hs.239147	guanine deaminase	92.31	28.00
20	430337	M36707	Hs.239600	calmodulin-like 3	1.18	1.08
	430378	Z29572	Hs.2556	tumor necrosis factor receptor superfam	5.28	66.00
	430388	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	16.76	38.00
	430393	BE185030	Hs.241305	estrogen-responsive B box protein	1.63	1.50
	430439	AL133561		DKFZP434B061 protein	1.00	1.00
25	430451	AA836472	Hs.297939	cathepsin B	1.64	2.12
	430454	AW469011	Hs.105635	ESTs	63.35	44.00
	430466	AF052573	Hs.241517	polymerase (DNA directed), theta	2.47	1.91
	430481	AA479678	Hs.203269	ESTs, Moderately similar to ALU8_HUMAN A	1.00	31.00
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
30	430508	AI015435	Hs.104637	ESTs	4.75	7.27
	430533	AA480895	Hs.57749	ESTs, Weakly similar to T17288 hypotheti	1.00	1.00
	430563	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.00	1.59
	430577	Z26317	Hs.94560	desmoglein 2	1.72	1.30
	430678	AA401369	Hs.190721	ESTs	0.90	17.00
35	430686	NM_001942	Hs.2633	desmoglein 1	1.00	1.00
	430788	AI742925	Hs.7179	ESTs, Weakly similar to 2004399A chromos	1.62	1.84
	430890	X54232	Hs.2699	glypican 1	1.58	1.40
	430935	AW072916		zinc finger protein 131 (clone pHZ-10)	90.28	132.00
	430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
40	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
	431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
	431092	AI332764	Hs.125757	ESTs	13.46	63.00
	431124	AF284221	Hs.59506	doublesex and mab-3 related transcriptio	49.43	62.00
	431164	AA493650	Hs.94367	Homo sapiens cDNA: FLJ23494 fis, clone L	0.44	2.20
45	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	182.26	101.00
	431221	AW207837	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	4.15	13.97
	431277	AA501806	Hs.345824	ESTs	1.00	86.00
	431322	AW970622		gb:EST382704 MAGE resequences, MAGK Homo	40.55	200.00
	431342	AW971018	Hs.21659	ESTs	1.00	53.00
50	431384	BE158000	Hs.285026	gb:MR2-HT0377-150200-202-e03 HT0377 Homo	0.94	1.14
	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	1.30	1.25
	431494	AA991355	Hs.298312	hypothetical protein DKFZp434A1315	3.90	26.00
	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	1.41	1.87
	431548	AI834273	Hs.9711	novel protein	5.66	15.00
55	431630	NM_002204	Hs.265829	integrin, alpha 3 (antigen CD49C, alpha	0.99	1.44
	431745	AW972448	Hs.163425	ESTs	0.99	3.51
	431770	BE221880	Hs.268555	5'-3' exoribonuclease 2	67.12	91.00
	431830	Y16645	Hs.271387	small inducible cytokine subfamily A (Cy	3.36	4.71
	431846	BE019924	Hs.271580	uroplakin 1B	4.49	2.51
60	431890	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	2.20	3.32
	431934	AB031481	Hs.272214	STG protein	1.01	1.04
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	51.17	46.35
	432006	AL137382	Hs.272320	Homo sapiens mRNA; cDNA DKFZp434L1226 (f	0.94	1.65
	432023	R43020	Hs.236223	EST	0.94	47.00
65	432201	AI538613	Hs.298241	Transmembrane protease, serine 3	1.10	2.24
	432210	AI567421	Hs.273330	Homo sapiens, clone IMAGE:3544662, mRNA,	1.42	1.45
	432226	AW182766	Hs.273558	phosphate cytidyltransferase 1, cholin	1.00	1.00
	432239	X81334	Hs.2936	matrix metalloproteinase 13 (collagenase	18.67	1.00
	432265	BE382679	Hs.285753	SCG10-like-protein	1.09	1.21
70	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	40.98	58.00
	432365	AK001106	Hs.274419	hypothetical protein FLJ10244	1.00	214.00
	432374	W68815	Hs.301885	Homo sapiens cDNA FLJ11346 fis, clone PL	157.34	37.00
	432375	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
	432407	AA221036		gb:zr03f12.r1 Stratagene NT2 neuronal pr	73.71	75.00
75	432441	AW292425	Hs.163484	ESTs	56.35	72.00
	432489	AI804855	Hs.207530	ESTs	1.00	24.00
	432543	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 fis, clone C	137.72	98.00
	432552	AI537170	Hs.173725	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.00	31.00
	432583	AW023624	Hs.162282	potassium channel TASK-4; potassium chan	0.27	35.18
80	432606	NM_002104	Hs.3066	granzyme K (serine protease, granzyme 3;	2.87	6.22
	432625	AI243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin	26.63	56.00
	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	1.92	5.29
	432677	NM_004482	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	45.13	31.00
85	432753	NM_014075	Hs.336938	Homo sapiens PRO0593 mRNA, complete cds	1.00	68.00
	432788	AA521091	Hs.178499	Homo sapiens cDNA: FLJ23117 fis, clone L	2.69	3.67
	432842	AW674093	Hs.334822	hypothetical protein MGC4485	1.22	1.34
	432867	AW016936	Hs.233364	ESTs	1.00	1.00
	432917	NM_014125	Hs.241517	PRO0327 protein	10.25	6.62



	432920	U37689	Hs.3128	polymerase (RNA) II (DNA directed) polyp	1.44	1.30
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	154.79	85.64
	433023	AW864793	Hs.87409	thrombospondin 1	20.96	100.00
5	433042	AW193534	Hs.281895	Homo sapiens cDNA FLJ11660 fis, clone HE	1.00	10.00
	433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
	433159	AB035898	Hs.150587	kinesin-like protein 2	13.82	39.00
	433183	AF231338	Hs.222024	transcription factor BMAL2	1.00	69.00
	433258	AA622788	Hs.203613	ESTs, Weakly similar to ALUB_HUMAN !!!!	1.00	1.25
10	433409	AI278802	Hs.25661	ESTs	44.81	117.00
	433437	U20536	Hs.3280	caspase 6, apoptosis-related cysteine pr	70.39	105.00
	433485	AI493076	Hs.201967	aldo-keto reductase family 1, member C2	11.55	2.00
	433537	AI733692	Hs.112488	ESTs	8.66	55.00
	433547	W04978	Hs.303023	beta tubulin 1, class VI	25.16	83.00
15	433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
	433647	AA603367	Hs.222294	ESTs	20.30	49.00
	433658	L03678	Hs.156110	immunoglobulin kappa constant	5.92	10.03
	433800	AI094221	Hs.135150	lung type-I cell membrane-associated gly	2.29	2.22
	433819	AW511097	Hs.112765	ESTs	3.71	8.00
20	433862	D86960	Hs.3610	KIAA0205 gene product	62.08	104.00
	433980	AA137152	Hs.286049	phosphoserine aminotransferase	108.91	47.00
	434088	AF116677	Hs.249270	hypothetical protein PRO1966	1.00	1.00
	434094	AA305599	Hs.238205	hypothetical protein PRO2013	121.27	87.00
	434105	AW952124	Hs.13094	presenilins associated rhomboid-like pro	1.22	1.23
25	434217	AW014795	Hs.23349	ESTs	14.11	57.00
	434340	AI193043	Hs.128685	ESTs, Weakly similar to T17226 hypotheti	2.10	2.56
	434360	AA401369	Hs.190721	ESTs	40.98	17.00
	434414	AI798376		gb:tr34b07.x1 NCI_CGAP_Ov23 Homo sapiens	1.48	1.56
	434424	AI811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
30	434467	BE552368	Hs.231853	Homo sapiens cDNA FLJ13445 fis, clone PL	54.91	85.00
	434551	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	2.46	2.00
	434627	AI221894	Hs.39311	ESTs	1.00	1.00
	434699	AA643687	Hs.149425	Homo sapiens cDNA FLJ11980 fis, clone HE	1.00	23.00
	434769	AA648884	Hs.134278	Homo sapiens cDNA FLJ12676 fis, clone NT	7.08	56.00
35	434792	AA649253	Hs.132458	ESTs	8.52	44.00
	434808	AF155108	Hs.256150	Homo sapiens, Similar to RIKEN cDNA 2810	11.33	1.00
	434828	D90070	Hs.96	phorbol-12-myristate-13-acetate-induced	1.00	1.00
	434876	AF160477	Hs.61460	Ig superfamily receptor LNIR	1.25	1.29
	434891	AA814309	Hs.123583	ESTs	1.00	6.00
40	434928	AW015595	Hs.4267	Homo sapiens clones 24714 and 24715 mRNA	1.00	1.00
	435013	H91923	Hs.110024	Target CAT	1.26	1.10
	435066	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin	1.69	1.37
	435087	AW975241	Hs.23567	ESTs	1.00	1.00
	435099	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.90	1.93
45	435159	AA668879	Hs.116649	ESTs	1.00	1.00
	435205	X54136	Hs.181125	immunoglobulin lambda locus	1.02	1.46
	435232	NM_001262	Hs.4854	cyclin-dependent kinase inhibitor 2C (p1	2.04	2.70
	435304	H10709	Hs.269524	ESTs	27.58	139.00
	435313	AI769400	Hs.189729	ESTs	1.00	14.00
50	435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
	435509	AI458679	Hs.181915	ESTs	1.00	1.00
	435525	AI831297	Hs.123310	ESTs	1.00	56.00
	435532	AW291488	Hs.117305	Homo sapiens, clone IMAGE:3682908, mRNA	1.00	2.00
	435550	AI224456	Hs.324507	H.sapiens polyA site DNA	3.42	3.92
55	435602	AF217515	Hs.283532	uncharacterized bone marrow protein BM03	3.95	1.80
	435766	R11673	Hs.186498	ESTs	1.00	28.00
	435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
	436069	AI056879	Hs.263209	ESTs	1.00	58.00
	436170	AW450381	Hs.14529	ESTs	1.00	18.00
60	436211	AK001581	Hs.334828	hypothetical protein FLJ10719; KIAA1794	5.84	22.00
	436213	AA325512	Hs.71472	hypothetical protein FLJ10774; KIAA1709	1.42	1.27
	436217	T53925	Hs.107	fibrinogen-like 1	57.97	31.00
	436238	AK002163	Hs.301724	hypothetical protein FLJ11301	2.51	1.71
	436251	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)	2.33	1.64
65	436291	BE568452	Hs.344037	protein regulator of cytokinesis 1	108.99	52.00
	436302	AL355841	Hs.99330	hypothetical protein FLJ23588	0.75	2.81
	436396	AW992292	Hs.152213	wingless-type MMTV integration site fami	60.01	1.00
	436414	BE264633	Hs.143638	WD repeat domain 4	2.50	2.19
	436419	AI948626	Hs.171356	ESTs	0.95	1.33
70	436443	AW138211	Hs.128746	ESTs	1.12	9.26
	436474	AJ270693	Hs.199887	ESTs	1.00	1.00
	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.28	1.56
	436486	AA742221	Hs.120633	ESTs	1.00	19.00
	436511	AA721252	Hs.291502	ESTs	16.76	14.00
75	436553	X57809	Hs.181125	immunoglobulin lambda locus	1.08	1.74
	436557	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	19.20	9.75
	436608	AA628980		down syndrome critical region protein DS	33.92	25.00
	436667	AW025183	Hs.127680	ESTs	0.89	1.19
	436771	AW975687	Hs.292979	ESTs	1.00	10.00
80	436839	AA401369	Hs.190721	ESTs	1.00	17.00
	436887	AW953157	Hs.193235	hypothetical protein DKFZp547D155	1.06	1.15
	436944	AW268614	Hs.5840	ESTs	1.00	1.00
	436961	AW375974	Hs.156704	ESTs	25.13	25.00
	436972	AA284679	Hs.25640	claudin 3	1.59	1.46
85	437016	AU076916	Hs.5398	guanine monphosphate synthetase	2.35	1.78
	437044	AL035864	Hs.69517	cDNA for differentially expressed CO16 g	1.34	1.13



	437181	AI306615	Hs.125343	ESTs, Weakly similar to KIAA0758 protein	1.00	17.00
	437204	AL110216	Hs.22826	ESTs, Weakly similar to I55214 salivary	40.55	82.00
	437205	AL110232	Hs.279243	Homo sapiens mRNA; cDNA DKFZp564D2071 (f	1.00	112.00
5	437259	AI377755	Hs.120695	ESTs	1.00	205.00
	437270	R18087	Hs.323769	cisplatin resistance related protein CRR	1.56	1.54
	437271	AL137445	Hs.28846	Homo sapiens mRNA; cDNA DKFZp566O134 (fr	113.25	125.00
	437370	AL359567	Hs.161962	Homo sapiens mRNA; cDNA DKFZp547D023 (fr	1.82	4.57
	437390	AI125859	Hs.112607	ESTs	1.35	1.75
10	437412	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	3.58	3.20
	437435	AI306152	Hs.27027	hypothetical protein DKFZp762H1311	3.03	1.08
	437444	H46008	Hs.31518	ESTs	1.00	39.00
	437568	AI954795	Hs.156135	ESTs	1.00	19.00
	437623	D63880	Hs.5719	chromosome condensation-related SMC-asso	1.95	1.57
15	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	1.00	3.00
	437814	AI088192	Hs.135474	ESTs, Weakly similar to DDX9_HUMAN ATP-D	1.00	45.00
	437840	AA884836	Hs.292014	ESTs	1.07	1.78
	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 [H.sa	1.68	3.26
	437879	BE262082	Hs.5894	hypothetical protein FLJ10305	1.87	2.52
20	437915	AI637993	Hs.202312	Homo sapiens clone N11 NTera2D1 teratoca	74.05	35.00
	437916	BE566249	Hs.20999	hypothetical protein FLJ23142	23.15	89.00
	437937	AI917222	Hs.121655	ESTs	1.00	1.00
	437942	AI888256	Hs.307526	ESTs	12.28	31.00
25	438091	AW373062		nuclear receptor subfamily 1, group 1, m	1.53	10.85
	438113	AI467908	Hs.8882	ESTs	1.80	2.39
	438119	AW963217	Hs.203961	ESTs, Moderately similar to AF116721 89	22.67	36.90
	438274	AI918906	Hs.55080	ESTs	1.00	1.00
	438378	AW970529	Hs.86434	hypothetical protein FLJ21816	38.92	38.00
	438403	AA806607	Hs.292206	ESTs	1.00	1.00
30	438494	AA908678	Hs.130183	ESTs	2.05	80.00
	438546	AW297204	Hs.125811	ESTs	1.00	131.00
	438552	AJ245820	Hs.6314	type I transmembrane receptor (seizure-r	1.43	1.45
	438702	AI879064	Hs.54618	ESTs	1.00	34.00
	438724	AW612553	Hs.114670	Human DNA sequence from clone RP11-16L21	1.33	1.10
35	438746	AI885815	Hs.184727	Human melanoma-associated antigen p97 (m	2.42	1.59
	438779	NM_003787	Hs.6414	nucleolar protein 4	1.00	18.00
	438821	AA826425	Hs.192375	ESTs	2.03	2.57
	438885	AI886558	Hs.184987	ESTs	6.42	88.00
	438898	AA401369	Hs.190721	ESTs	22.41	17.00
40	438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	1.00	1.00
	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9	2.20	1.88
	439000	AW979121		gb:EST391231 MAGE resequences, MAGP Homo	2.78	4.81
	439023	AA745978	Hs.28273	ESTs	1.17	1.31
	439024	R96696	Hs.35598	ESTs	1.00	28.00
45	439128	AI949371	Hs.153089	ESTs	1.00	67.00
	439146	AW138909	Hs.156110	immunoglobulin kappa constant	1.38	1.41
	439223	AW238299	Hs.250618	UL16 binding protein 2	1.93	1.64
	439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
	439318	AW837046	Hs.6527	G protein-coupled receptor 56	2.00	2.20
50	439343	AF086161	Hs.114611	hypothetical protein FLJ11808	6.10	7.37
	439394	AA401369	Hs.190721	ESTs	3.39	17.00
	439410	AA632012	Hs.188746	ESTs	1.83	3.07
	439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00
	439452	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	18.76	122.00
55	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13	2.78	1.58
	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	1.22	1.44
	439492	AF086310	Hs.103159	ESTs	7.46	39.00
	439523	W72348	Hs.185029	ESTs	1.00	1.19
	439592	AF086413	Hs.58399	ESTs	1.00	1.00
60	439606	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
	439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
	439702	AW085525	Hs.134182	ESTs	4.30	10.00
	439706	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	86.55	11.00
	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	2.36	1.88
65	439750	AL359053	Hs.57664	Homo sapiens mRNA full length insert cDN	2.02	6.08
	439759	AL359055	Hs.67709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	439780	AL109688		gb:Homo sapiens mRNA full length insert	7.27	25.00
	439840	AW449211	Hs.105445	GDNF family receptor alpha 1	1.00	1.00
	439926	AW014875	Hs.137007	ESTs	32.58	71.00
70	439963	AW247529	Hs.6793	platelet-activating factor acetylhydrola	21.28	9.55
	439979	AW600291	Hs.6823	hypothetical protein FLJ10430	68.83	61.00
	440006	AK000517	Hs.6844	hypothetical protein FLJ20510	1.83	4.02
	440028	AW473675	Hs.125843	ESTs, Weakly similar to T17227 hypotheti	1.42	2.54
	440106	AA864968	Hs.127699	KIAA1603 protein	1.00	54.00
75	440138	AB033023	Hs.318127	hypothetical protein FLJ10201	24.18	52.00
	440273	AI805392	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	3.21	4.72
	440289	AW450991	Hs.192071	ESTs	38.63	113.00
	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
	440492	R39127	Hs.21433	hypothetical protein DKFZp547J036	2.35	3.62
80	440527	AV657117	Hs.184164	ESTs, Moderately similar to S65657 alpha	10.84	57.00
	440659	AF134160	Hs.7327	claudin 1	3.18	2.37
	440704	M69241	Hs.162	insulin-like growth factor binding prote	2.89	2.09
	440943	AW082298	Hs.146161	hypothetical protein MGC2408	2.02	1.41
	440994	AI160011	Hs.272068	ESTs	1.29	1.14
85	441020	AA401369	Hs.190721	ESTs	142.99	17.00
	441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00



	441128	AA570256		ESTs, Weakly similar to T23273 hypotheti	4.13	3.50
	441290	W27501	Hs.89605	cholinergic receptor, nicotinic, alpha p	1.00	1.00
	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	130.23	43.00
5	441377	BE218239	Hs.202656	ESTs	22.03	1.00
	441390	AI692560	Hs.131175	ESTs	3.65	7.70
	441497	R51064	Hs.23172	ESTs	1.00	1.00
	441525	AW241867	Hs.127728	ESTs	1.53	1.42
	441553	AA281219	Hs.121296	ESTs	1.89	1.57
10	441607	NM_005010	Hs.7912	neuronal cell adhesion molecule	1.47	2.11
	441633	AW958544	Hs.112242	normal mucosa of esophagus specific 1	216.22	363.00
	441636	AA081846	Hs.7921	Homo sapiens mRNA; cDNA DKFZp566E183 (fr	2.31	2.05
	441737	X79449	Hs.7957	adenosine deaminase, RNA-specific	1.30	1.49
	441790	AA401369	Hs.190721	ESTs	44.15	17.00
	441801	AW242799	Hs.86366	ESTs	1.00	1.00
15	441919	AI553802	Hs.128121	ESTs	1.00	122.00
	441937	R41782	Hs.22279	ESTs	0.86	1.37
	441954	AI744935	Hs.8047	Fanconi anemia, complementation group G	1.48	1.39
	442025	AW887434	Hs.11810	CDA11 protein	1.00	46.00
	442029	AW956698	Hs.14456	neural precursor cell expressed, develop	9.92	45.00
20	442072	AI740832	Hs.12311	Homo sapiens clone 23570 mRNA sequence	25.05	77.00
	442108	AW452649	Hs.166314	ESTs	3.61	3.14
	442117	AW664964	Hs.128899	ESTs	3.00	5.49
	442137	AA977235	Hs.128830	ESTs, Weakly similar to Z192_HUMAN ZINC	1.00	1.00
25	442159	AW163390	Hs.278554	heterochromatin-like protein 1	1.92	1.66
	442179	AA983842	Hs.333555	chromosome 2 open reading frame 2	27.22	50.00
	442328	AI952430	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	5.00	3.42
	442432	BE093589	Hs.38178	hypothetical protein FLJ23468	181.59	76.00
	442530	AI580830	Hs.176508	Homo sapiens cDNA FLJ14712 fis, clone NT	10.59	144.00
30	442547	AA306997	Hs.217484	ESTs, Weakly similar to ALU1_HUMAN ALU S	109.23	98.00
	442556	AL137761	Hs.8379	Homo sapiens mRNA; cDNA DKFZp586L2424 (f	1.00	53.00
	442619	AA447492	Hs.20183	ESTs, Weakly similar to AF164793 1 prote	29.02	50.00
	442710	AI015631	Hs.23210	ESTs	1.00	19.00
	442717	R88362	Hs.180591	ESTs, Weakly similar to T23976 hypotheti	1.00	5.00
35	442875	BE623003	Hs.23625	Homo sapiens clone TCCCTA00142 mRNA sequ	22.85	50.00
	442914	AW188551	Hs.99519	hypothetical protein FLJ14007	25.33	82.00
	442932	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	3.18	4.41
	442942	AW167087	Hs.131562	ESTs	8.45	64.00
	443068	AI188710		ESTs	1.00	27.00
40	443204	AW205878	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	1.00	24.00
	443211	AI128388	Hs.143655	ESTs	12.42	2.00
	443247	BE614387	Hs.333893	c-Myc target JPO1	128.84	96.00
	443324	R44013	Hs.164225	ESTs	0.02	4.59
	443383	AI792453	Hs.166507	ESTs	1.00	47.00
45	443400	R28424	Hs.250648	ESTs	18.52	61.00
	443426	AF098158	Hs.9329	chromosome 20 open reading frame 1	4.02	1.75
	443572	AA025610	Hs.9605	cleavage and polyadenylation specific fa	2.98	2.57
	443575	AI078022	Hs.269636	ESTs, Weakly similar to ALU1_HUMAN ALU S	1.00	29.00
	443614	AV655386	Hs.7645	fibrinogen, B beta polypeptide	1.00	16.00
50	443633	AL031290	Hs.9654	similar to pregnancy-associated plasma p	1.00	39.00
	443648	AI085377	Hs.143610	ESTs	39.81	70.00
	443715	AI583187	Hs.9700	cyclin E1	48.74	7.00
	443723	AI144442	Hs.157144	syntaxin 6	1.29	1.30
	443802	AW504924	Hs.9805	KIAA1291 protein	1.75	1.61
55	443859	NM_013409	Hs.9914	folistatin	1.35	1.13
	443892	AA401369	Hs.190721	ESTs	1.00	17.00
	443947	W24187		gb:zb47f09.r1 Soares_fetal_lung_NbHL19W	1.33	1.64
	443991	NM_002250	Hs.10082	potassium intermediate/small conductance	5.71	6.87
	444006	BE395085	Hs.10086	type I transmembrane protein Fn14	1.47	1.92
60	444009	AI380792	Hs.135104	ESTs	1.00	77.00
	444017	U04840	Hs.214	neuro-oncological ventral antigen 1	1.00	1.00
	444127	N63620	Hs.13281	ESTs	1.00	29.00
	444129	AW294292	Hs.256212	ESTs	1.00	1.00
	444279	U62432	Hs.89605	cholinergic receptor, nicotinic, alpha p	0.60	7.80
65	444371	BE540274	Hs.239	forkhead box M1	2.91	1.14
	444378	R41339	Hs.12569	ESTs	1.00	1.00
	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti	469.00	556.00
	444461	R53734	Hs.25978	ESTs, Weakly similar to 2109260A B cell	12.88	105.00
	444471	AB020684	Hs.11217	KIAA0877 protein	24.91	90.00
	444489	AI151010	Hs.157774	ESTs	1.00	111.00
70	444619	BE538082	Hs.8172	ESTs, Moderately similar to A46010 X-lin	1.00	70.00
	444665	BE613126	Hs.47783	B aggressive lymphoma gene	30.56	139.00
	444707	AI188613	Hs.41690	desmocollin 3	1.00	1.00
	444735	BE019923	Hs.243122	hypothetical protein FLJ13057 similar to	77.02	90.00
75	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	1.57	1.31
	444783	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	77.55	2.00
	445236	AK001676	Hs.12457	hypothetical protein FLJ10814	1.00	27.00
	445258	AI635931	Hs.147613	ESTs	1.00	73.00
	445413	AA151342	Hs.12677	CGI-147 protein	28.14	50.00
80	445417	AK001058	Hs.12680	Homo sapiens cDNA FLJ10196 fis, clone HE	1.81	2.62
	445443	AV653838	Hs.322971	ESTs	1.00	1.00
	445462	AA378776	Hs.288649	hypothetical protein MGC3077	2.09	1.70
	445517	AF208855	Hs.12830	hypothetical protein	1.87	70.00
	445537	AJ245671	Hs.12844	EGF-like-domain, multiple 6	1.71	2.72
85	445580	AF167572	Hs.12912	skb1 (S. pombe) homolog	1.52	1.34
	445654	X91247	Hs.13046	thioredoxin reductase 1	1.51	1.52



	445669	AI570830	Hs.174870	ESTs	10.95	11.45
	445818	BE045321	Hs.136017	ESTs	1.00	1.00
	445873	AA250970	Hs.251946	poly(A)-binding protein, cytoplasmic 1-l	49.42	54.00
5	445885	AI734009	Hs.127699	KIAA1603 protein	1.00	132.00
	445898	AF070623	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.00	1.00
	445903	AI347487	Hs.132781	class I cytokine receptor	1.00	36.00
	445932	BE046441	Hs.333555	Homo sapiens clone 24859 mRNA sequence	2.41	2.88
	445982	BE410233	Hs.13501	pescadillo (zebrafish) homolog 1, contai	1.60	1.35
10	446078	AI339982	Hs.156061	ESTs	1.00	42.00
	446102	AW168067	Hs.317694	ESTs	1.00	1.00
	446157	BE270828	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	1.70	1.53
	446269	AW263155	Hs.14559	hypothetical protein FLJ10540	73.01	48.00
	446292	AF081497	Hs.279682	Rh type C glycoprotein	1.55	1.26
	446293	AI420213	Hs.149722	ESTs	1.00	2.00
15	446423	AW139655	Hs.150120	ESTs	1.10	4.19
	446428	AW082270	Hs.12496	ESTs, Weakly similar to ALU4_HUMAN ALU S	0.53	3.26
	446432	AI377320	Hs.150058	ESTs	1.00	5.00
	446528	AU076640	Hs.15243	nucleolar protein 1 (120kD)	1.36	1.31
20	446574	AI310135	Hs.335933	ESTs	3.89	72.00
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	32.03	20.23
	446636	AC002563	Hs.15767	citron (rho-interacting, serine/threonin	4.19	5.07
	446783	AW138343	Hs.141867	ESTs	2.82	9.47
	446839	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	110.28	28.00
25	446849	AU076617	Hs.16251	cleavage and polyadenylation specific fa	3.26	2.94
	446856	AI814373	Hs.164175	ESTs	6.38	11.30
	446872	X97058	Hs.16362	pyrimidinergic receptor P2Y, G-protein c	1.98	2.03
	446880	AI811807	Hs.108646	Homo sapiens cDNA FLJ14934 fis, clone PL	94.90	113.00
	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	1.67	3.90
	446989	AK001898	Hs.16740	hypothetical protein FLJ11036	2.82	3.12
30	447022	AW291223	Hs.157573	ESTs	1.00	170.00
	447033	AI357412	Hs.157601	ESTs	7.15	107.00
	447078	AW885727	Hs.9914	ESTs	47.24	24.00
	447081	Y13896	Hs.17287	potassium inwardly-rectifying channel, s	0.12	17.88
35	447131	NM_004585	Hs.17466	retinoic acid receptor responder (tazaro	0.97	1.48
	447149	BE298857	Hs.326	TAR (HIV) RNA-binding protein 2	1.24	1.26
	447153	AA805202	Hs.315562	ESTs	1.00	54.00
	447164	AF026941	Hs.17518	Homo sapiens cig5 mRNA, partial sequence	1.00	67.00
	447178	AW594641	Hs.192417	ESTs	3.42	50.00
40	447250	AI878909	Hs.17883	protein phosphatase 1G (formerly 2C), ma	1.60	1.52
	447289	AW247017	Hs.36978	melanoma antigen, family A, 3	1.00	1.00
	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m	146.62	51.00
	447350	AI375572	Hs.172634	ESTs	1.00	12.00
45	447377	N27687	Hs.334334	transcription factor AP-2 alpha (activat	2.55	63.00
	447415	AW937335	Hs.28149	ESTs, Weakly similar to KF3B_HUMAN KINES	0.91	1.13
	447425	AI963747	Hs.18573	acylphosphatase 1, erythrocyte (common)	1.00	35.00
	447519	U46258	Hs.339665	ESTs	59.89	49.00
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
50	447534	AA401369	Hs.190721	ESTs	1.00	17.00
	447636	Y10043		high-mobility group (nonhistone chromoso	1.41	1.11
	447688	N87079	Hs.19236	Target CAT	1.00	39.00
	447733	AF157482	Hs.19400	MAD2 (mitotic arrest deficient, yeast, h	1.17	1.12
	447769	AW873704	Hs.320831	Homo sapiens cDNA FLJ14597 fis, clone NT	6.47	5.95
55	447802	AW593432	Hs.161455	ESTs	0.73	2.34
	447850	AB018298	Hs.19822	SEC24 (S. cerevisiae) related gene famil	86.45	116.00
	447924	AI817226	Hs.313413	ESTs, Weakly similar to T23110 hypotheti	1.00	1.00
	447973	AB011169	Hs.20141	similar to S. cerevisiae SSM4	3.50	4.27
	448030	N30714	Hs.325960	membrane-spanning 4-domains, subfamily A	4.13	142.00
60	448105	AI538613	Hs.298241	Transmembrane protease, serine 3	1.15	2.24
	448243	AW369771	Hs.52620	integrin, beta 8	15.84	1.00
	448278	W07369	Hs.11782	ESTs	0.97	1.90
	448290	AK002107	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	1.00	1.00
	448296	BE622756	Hs.10949	Homo sapiens cDNA FLJ14162 fis, clone NT	2.42	2.17
65	448357	BE274396	Hs.108923	RAB38, member RAS oncogene family	1.44	1.08
	448390	AL035414	Hs.21068	hypothetical protein	1.00	43.00
	448469	AW504732	Hs.21275	hypothetical protein FLJ11011	2.63	2.49
	448569	BE382657	Hs.21486	signal transducer and activator of trans	1.84	2.53
	448663	BE614599	Hs.106823	hypothetical protein MGC14797	3.29	46.00
70	448672	AI955511	Hs.225106	ESTs	1.00	21.00
	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	1.82	1.08
	448741	BE614567	Hs.19574	hypothetical protein MGC5469	2.48	1.92
	448757	AI366784	Hs.48820	TATA box binding protein (TBP)-associate	23.53	20.00
	448775	AB025237	Hs.388	nudix (nucleoside diphosphate linked moi	2.34	1.97
75	448826	AI580252	Hs.293246	ESTs, Weakly similar to putative p150 [H	74.07	62.67
	448830	AL031658	Hs.22181	hypothetical protein dJ310013.3	1.37	1.31
	448844	AI581519	Hs.177164	ESTs	1.00	31.00
	448988	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.84	1.95
	448993	AI471630		KIAA0144 gene product	1.63	1.49
80	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
	449029	N28989	Hs.22891	solute carrier family 7 (cationic amino	1.97	2.26
	449040	AF040704	Hs.149443	putative tumor suppressor	0.97	1.56
	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	27.13	90.00
	449053	AI625777	Hs.344766	ESTs	8.33	44.00
85	449054	AF148848	Hs.22934	myoneurin	73.85	104.00
	449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00



	449167	TO5095	Hs.19597	KIAA1694 protein	1.61	2.36
	449207	AL044222	Hs.23255	nucleoporin 155kD	2.36	1.56
	449228	AJ403107	Hs.148590	protein related with psoriasis	1.15	1.15
5	449230	BE613348	Hs.211579	melanoma cell adhesion molecule	206.65	151.00
	449305	AI638293		gb:tt09b07.x1 NCI_CGAP_GC6 Homo sapiens	17.28	45.00
	449318	AW236021	Hs.78531	Homo sapiens, Similar to RIKEN cDNA 5730	26.39	35.00
	449448	D60730	Hs.57471	ESTs	1.00	1.00
	449467	AW205006	Hs.197042	ESTs	1.00	1.00
10	449523	NM_000579	Hs.54443	chemokine (C-C motif) receptor 5	56.80	216.86
	449722	BE280074	Hs.23960	cyclin B1	150.03	1.00
	449976	H06350	Hs.135056	Human DNA sequence from clone RP5-850E9	2.16	2.85
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitter)	1.17	1.45
	450098	W27249	Hs.8109	hypothetical protein FLJ21080	1.79	2.38
15	450101	AV649989	Hs.24385	Human hbc647 mRNA sequence	1.00	69.00
	450149	AW969781	Hs.132863	Zic family member 2 (odd-paired Drosophila)	1.00	1.00
	450193	AI916071	Hs.15607	Homo sapiens Fanconi anemia complementation	29.85	34.00
	450221	AA328102	Hs.24641	cytoskeleton associated protein 2	1.00	1.00
	450372	BE218107	Hs.202436	ESTs	1.00	1.00
20	450375	AA009647	Hs.8850	a disintegrin and metalloproteinase domain	51.26	93.00
	450447	AF212223	Hs.25010	hypothetical protein P15-2	123.20	181.00
	450568	AL050078	Hs.25159	Homo sapiens cDNA FLJ10784 fis, clone NT	1.00	19.00
	450589	AI701505	Hs.202526	ESTs	1.00	23.00
	450684	AA872605	Hs.25333	interleukin 1 receptor, type II	1.00	100.00
25	450701	H39960	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.89	1.55
	450705	U90304	Hs.25351	iroquois homeobox protein 2A (IRX-2A) (	1.00	45.00
	450832	AA401369	Hs.190721	ESTs	25.17	17.00
	450937	R49131	Hs.26267	ATP-dependant interferon response protein	90.92	90.00
	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	3.33	1.70
30	451105	AI761324		gb:wi60b11.x1 NCI_CGAP_Co16 Homo sapiens	15.02	124.00
	451110	AI955040	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	143.00
	451253	H48299	Hs.26126	claudin 10	3.02	2.29
	451291	R39288	Hs.6702	ESTs	1.00	1.00
	451320	AW498974		diacylglycerol kinase, zeta (104kD)	2.92	18.00
35	451380	H09280	Hs.13234	ESTs	6.90	6.67
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant)	35.75	72.00
	451437	H24143	Hs.31945	hypothetical protein FLJ11071	1.00	69.00
	451462	AK000367	Hs.26434	hypothetical protein FLJ20360	1.83	2.10
	451524	AK001466	Hs.26516	hypothetical protein FLJ10604	1.13	1.07
40	451541	BE279383	Hs.26557	plakophilin 3	1.88	1.33
	451592	AI805416	Hs.213897	ESTs	1.00	1.00
	451635	AA018899	Hs.127179	cryptic gene	1.52	1.92
	451743	AA401369	Hs.190721	ESTs	4.95	17.00
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	13.55	31.00
45	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	451871	AI821005	Hs.118599	ESTs	1.81	2.53
	451952	AL120173	Hs.301663	ESTs	1.00	22.00
	452012	AA307703	Hs.279766	kinesin family member 4A	3.43	2.26
	452046	AB018345	Hs.27657	KIAA0802 protein	56.59	19.00
50	452194	AI694413	Hs.332649	olfactory receptor, family 2, subfamily	1.67	4.09
	452206	AW340281	Hs.33074	Homo sapiens, clone IMAGE:3606519, mRNA,	9.31	53.00
	452240	AA401369	Hs.190721	ESTs	13.42	17.00
	452256	AK000933	Hs.28661	Homo sapiens cDNA FLJ10071 fis, clone HE	39.03	94.00
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	153.01	340.00
55	452291	AF015592	Hs.28853	CDC7 (cell division cycle 7, S. cerevisiae)	1.95	23.00
	452295	BE379936	Hs.28866	programmed cell death 10	42.33	61.00
	452304	AA025386	Hs.61311	ESTs, Weakly similar to S10590 cysteine	1.17	2.14
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodomain	1.00	13.00
	452349	AB028944	Hs.29189	ATPase, Class VI, type 11A	1.09	1.42
60	452367	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	54.49	53.00
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced protein	1.00	32.00
	452410	AL133619		Homo sapiens mRNA; cDNA DKFZp434E2321 (f	1.26	1.99
	452461	N78223	Hs.108106	transcription factor	24.47	35.00
	452571	W31518	Hs.34665	ESTs	54.61	102.00
65	452613	AA461599	Hs.23459	ESTs	1.39	1.32
	452699	AW295390	Hs.213062	ESTs	1.00	26.00
	452705	H49805	Hs.246005	ESTs	1.00	1.00
	452747	AF160477	Hs.61460	Ig superfamily receptor LNIR	112.87	1.29
	452787	AW294022	Hs.222707	KIAA1718 protein	1.00	1.00
70	452795	AW392555	Hs.18878	hypothetical protein FLJ21620	1.00	1.00
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	7.91	75.00
	452833	BE559681	Hs.30736	KIAA0124 protein	3.16	1.92
	452838	U65011	Hs.30743	preferentially expressed antigen in melanoma	174.35	1.00
	452862	AA401369	Hs.190721	ESTs	98.26	17.00
75	452865	AW173720	Hs.345805	ESTs, Weakly similar to A47582 B-cell granule	1.55	1.00
	452934	AA581322	Hs.4213	hypothetical protein MGC16207	1.73	1.19
	452946	X95425	Hs.31092	EphA5	1.00	1.00
	452976	R44214	Hs.101189	ESTs	1.58	1.98
	453028	AB006532	Hs.31442	RecQ protein-like 4	1.80	1.60
80	453095	AW295660	Hs.252756	ESTs	0.77	1.50
	453102	NM_007197	Hs.31664	frizzled (Drosophila) homolog 10	1.00	1.00
	453103	AI301052	Hs.153444	ESTs	1.00	1.00
	453120	AA292891	Hs.31773	pregnancy-induced growth inhibitor	1.23	1.20
	453153	N53893	Hs.24360	ESTs	1.00	83.00
	453160	AI263307	Hs.239884	H2B histone family, member L	1.00	30.00
85	453197	AI916269	Hs.109057	ESTs, Weakly similar to ALU5_HUMAN ALU S	1.00	134.00



5	453210	AL133161	Hs.32360	hypothetical protein FLJ10867	1.69	1.93		
	453240	AI969564	Hs.166254	hypothetical protein DKFZp566i133	1.00	1.00		
	453317	NM_002277	Hs.41696	keratin, hair, acidic,1	1.19	1.27		
	453323	AF034102	Hs.32951	solute carrier family 29 (nucleoside tra	4.90	4.11		
	453331	AI240665	Hs.8860	ESTs	199.42	340.00		
	453392	U23752	Hs.32964	SRV (sex determining region Y)-box 11	1.00	16.00		
	453431	AF094754	Hs.32973	glycine receptor, beta	1.00	1.00		
	453439	AI572438	Hs.32976	guanine nucleotide binding protein 4	3.44	5.17		
10	453459	BE047032	Hs.257789	ESTs	2.84	5.58		
	453563	AW608906.comp	Hs.181163	hypothetical protein MGC5629	4.58	90.00		
	453633	AA357001	Hs.34045	hypothetical protein FLJ20764	1.74	1.60		
	453775	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	19.49	1.00		
	453830	AA534296	Hs.20953	ESTs	24.92	25.00		
15	453857	AL080235	Hs.35861	DKFZP586E1621 protein	167.59	66.00		
	453867	AI929383	Hs.33032	hypothetical protein DKFZp434N185	1.00	39.00		
	453883	AI638516	Hs.347524	cofactor required for Sp1 transcriptiona	1.97	1.58		
	453884	AA355925	Hs.36232	KIAA0186 gene product	63.89	20.00		
	453900	AW003582	Hs.226414	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.41	16.00		
20	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1	7.09	22.00		
	453941	U39817	Hs.36820	Bloom syndrome	29.75	19.00		
	453964	AI961486	Hs.12744	ESTs	1.00	1.00		
	453968	AA847843	Hs.62711	Homo sapiens, clone IMAGE:3351295, mRNA	2.06	1.81		
	453976	BE463830	Hs.163714	ESTs	3.02	131.00		
25	454024	AA993527	Hs.293907	hypothetical protein FLJ23403	1.00	131.00		
	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member	1.23	1.02		
	454042	T19228	Hs.172572	hypothetical protein FLJ20093	30.63	171.00		
	454059	NM_003154	Hs.37048	statherin	1.00	1.00		
	454066	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid	1.01	1.45		
30	454098	W27953	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11		
	454241	BE144666		gb:CM2-HT0176-041099-017-c02 HT0176 Homo	6.33	5.04		
	454417	AI244459	Hs.110826	trinucleotide repeat containing 9	4.30	7.82		
	454439	AW819152	Hs.154320	DKFZP566O1646 protein	1.00	1.00		
	455175	AW993247		gb:RC2-BN0033-180200-014-h09 BN0033 Homo	13.75	103.00		
35	455601	AI368680	Hs.816	SRV (sex determining region Y)-box 2	206.11	1.00		
	456237	AA203682		gb:zx52e07.r1 Soares_fetal_liver_spleen_	1.00	1.00		
	456321	NM_001327	Hs.87225	cancer/testis antigen	1.14	1.10		
	456475	NM_000144	Hs.95998	Friedreich ataxia	1.00	48.00		
	456508	AA502764	Hs.123469	ESTs, Weakly similar to AF208855 1 BM-01	162.25	189.00		
40	456534	X91195	Hs.100623	phospholipase C, beta 3, neighbor pseudo	2.12	1.80		
	456736	AW248217	Hs.1619	achaete-scute complex (Drosophila) homol	1.15	1.94		
	456759	BE259150	Hs.127792	delta (Drosophila)-like 3	1.00	1.00		
	456990	NM_004504	Hs.171545	HIV-1 Rev binding protein	16.42	84.00		
	457200	U33749	Hs.197764	thyroid transcription factor 1	0.57	1.76		
45	457234	AW968360	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	2.71	4.15		
	457465	AW301344	Hs.122908	DNA replication factor	46.37	47.00		
	457489	AI693815	Hs.127179	cryptic gene	1.12	1.35		
	457646	AA725650	Hs.112948	ESTs	1.55	2.51		
	457733	AW974812	Hs.291971	ESTs	1.00	55.00		
50	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein	4.36	3.18		
	458092	BE545684	Hs.343566	KIAA0251 protein	1.00	1.32		
	458098	BE550224		metallothionein 1E (functional)	1.00	22.00		
	458207	T28472	Hs.7655	U2 small nuclear ribonucleoprotein auxil	2.06	1.88		
	458242	BE299588	Hs.28465	Homo sapiens cDNA: FLJ21869 fis, clone H	1.00	1.00		
55	458247	R14439	Hs.209194	ESTs	7.00	9.85		
	458679	AW975460	Hs.142913	ESTs	1.00	3.00		
	458778	AW451034	Hs.326525	arylsulfatase D	1.31	2.01		
	458933	AI638429	Hs.24763	RAN binding protein 1	1.98	1.71		
	459352	AW810383	Hs.206828	ESTs	12.60	63.00		
60	459670	F01020	Hs.172004	titin	1.00	1.00		
	459702	AI204995		gb:an03c03.x1 Stratagene schizo brain S1	1.00	237.00		

TABLE 9B

65	Pkey:	Unique Eos probeset identifier number	
	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
70	Pkey	CAT Number	Accession
	407746	10125_1	AK001962 R69415 BE464605 AA418699 AA053293 AA149075 AA058396 AW338226 AW272659 AA454607 AI139535 AW469852 AI275461 AW271982 AA730033 AA576507 AA991217 AA782067 AI985851 AA805864 AA505598 AW469857 R69546 AA988279 AW001647 N63320 D82661 T27343 AA306950 AA360989 R58778
	408070	1036688_1	AW148852 BE350895
75	408660	107294_1	AA525775 AA056342 AI538978 AW975281 AA664986
	409522	113735_1	AA075382 AA075431
	409866	1156522_1	AW502152 H41202 H29772
	410032	1170435_1	BE065985 BE065944 BE066008 BE066083 BE066093
	411089	123172_1	AA456454 AA713730 AA091294 AA584921 N86077 AW836781 AA601031 AA579876 AA551106 AA633188 AW905577 AI955808 AI679386 AI679895 AA514764 AA454562 AI082382 AA595822 AA551351 AA586369 AA666384 AA188934 AA666398 AA551297 AA565188
80	411152	1234028_1	BE069199 AW936012 AW877466 AW819782 AW935798 AW835546 AW936042 BE069121 AW835625 AW877536 AW935885 BE069202 AW820019 AW935937 BE160180 AW935946 BE069101 BE069125 AW877527 BE160316 BE160398 AW935794 AW835701 AW935784 AL031778 X59711 NM_002505 M59079 AI870439 AI494259 AW664010 AA405063 AA436132 BE174516 AA412691 AI400314 AA436024 T29403 BE079412 BE079428 N90322 AI631202 AI141758 AI016793 AI167566 AI862075 AI375230 AI208445 AW235763 AL044113 AA382556 AW953918 AA927051 AA889823 BE003094 AW390155 AW360805 AW360823 AW360810 AA425472 AI694282 AL044114 AI684577 AI809865
	412537	1304_1	



			AI478773 AI160445 AI674630 N69088 AW665529 N49278 AI129239 AI457890 AI621264 AW297152 AI268215 AA907787 AI286170 AI017982 AI963541 AI469807 AI969353 BE552356 N66509 AA736741 AA382555 AW075811 AW292026 H06382 AW957730 AA352014 R13591 AA121201 D60420 BE263253 BE047862 Z41952 AI424991 AI693507 AI863108 AA599060 AI091148 AA598689 R39887 AA813482 AW016452 H06383 R41807 AI364268 AA620528 AI241940 AW089149 AW090733 AW088875 Z38240 AA121202 R17734 BE157489 BE157560 AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245 AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW960304 T29812 AA476873 BE297387 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 AI719075 BE270172 BE269819 AA889955 AI204630 W25243 AI935150 AA872039 W72395 T99630 AI422691 H98460 N31428 BE255916 H03265 AI857576 AA776920 AA910644 AA459522 AA293140 AW514667 R75953 AW662396 AA662522 AI865147 AI423153 AW262230 AA584410 AA583187 AW024595 AW069734 AI828996 AA282997 AA876046 AW613002 AA527373 AW972459 AI831360 AA621337 AA100926 AA772418 AA594628 AI033892 W95096 AI034317 AA398727 AI085031 N95210 AI459432 AI041437 AA932124 AA627684 AA935829 AI004827 AI423513 AI094597 H42079 R54703 AI630359 AA617681 AA978045 AA643280 W44561 AI991988 AI537692 AI090262 AA740817 AI312104 AI911822 AA416871 AI185409 AA129784 AA701623 AI075239 AI139549 AA633648 AI339996 AI336880 AA399239 AI078708 AI085351 AI362835 AI346618 AI146955 AI989380 AI348243 N92892 AA765850 AI494230 AI278887 AA962596 AI492600 W80435 AA001979 R97424 AI129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785 AI494211 AW059601 AW886710 R92790 N59755 AI361128 AW589407 H47725 H97534 H48076 H48450 T99631 AW300758 H03431 R76789 AA954344 H77576 R96823 AI457100 N92845 N49682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923 H03266 BE261919 AA769633 AA480310 AA507454 AA910586 AI203723 AW104725 W25611 W25071 T88980 H03513 T77589 R99156 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283672 AI267700 AI720344 AA191424 AI023543 AI469633 AA172056 AW958465 AA172236 AW953397 AA355086 AW265494 AA455904 AA195677 AW265432 AW991605 AA456370 N28754 N28747 AI568146 AI979339 AA322671 AA322672 AW955043 AI990326 AA776406 AI016250 AA843678 AW451882 N23137 N23129 W70051 AI038748 AA831327 AI925845 AW945895 Z42183 T31621 T97478 D62703 AA242966 D79798 AU076704 T74854 T74860 T72098 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72826 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M64982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 R06796 H48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AA011465 AA345378 AV654847 AV654272 AV656001 AI064740 T82897 N33594 AA344542 AW805054 AI207457 T61743 AA026737 H94389 AA382695 AA918409 T68044 S82092 T39959 AI017721 AA312395 AA312919 T40156 H66239 AV652989 H38728 R98521 AV655200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343596 AW470774 AV651256 N54417 AA812862 AW182929 AI111192 H61463 H72060 AA344503 H38639 AI277511 AV661108 AI207625 T47810 AA235252 T27853 T47778 R95746 H70620 AA701463 AW827166 R98475 C20925 AV657287 T71959 T71313 T73920 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T27854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 AI110639 AA344603 AF063513 T64696 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74677 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA342807 AA341964 T53747 T72042 T62764 AI064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 AI207557 T68219 W86031 T69081 T64232 R93196 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 AI248502 R29454 T64764 T57001 T73052 T71429 T51176 T58866 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 AI792788 BE142230 AA252019 AI910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 AI571948 AA507595 AA614579 AA587613 R83818 AA568312 AA614409 AA307578 AI925552 AW950155 AI910083 M12075 BE074052 AW004668 AA578674 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA858276 R55389 AI001051 AW050700 AW750216 AA614539 BE074045 AI307407 AW602303 BE073575 AI202532 AA524242 AI970839 AI909751 BE076078 AI909749 R55292 AW881145 AA490718 M85637 AA304575 T06067 AA331991 AL119930 AA320696 AW752565 AL031985 AL137241 AI792386 AI733664 AI857654 AI049911 AA337221 AA336756 AW966196 AW953120 R56325 AA349562 AI493134 AI498691 AW771508 AI498457 AI768408 AI783624 AI383985 AI580267 D79813 AA393768 AK001536 AA191092 AW510354 AI554256 AL353968 AA134266 AA663848 AA400100 AA401424 AL038843 AA161338 BE268213 AA425597 N87306 AA092969 BE566038 AA247451 N47392 AI928802 AW182584 AW027872 AI819831 AI936994 W56258 AI653448 AI278611 AI283557 AI824306 AW338658 AW150899 AA687514 N47393 N29885 AA973469 AI038904 AI292064 AI034339 AW674593 N72156 AI079733 AI038683 AI291616 AA491599 AA993675 AA837380 BE006554 BE006473 AI087090 T33044 AA652043 AI203503 AA583959 W35283 AI129926 Z41844 AW020925 AW575848 AI684603 AA493297 AI140689 AI277175 AA425444 AI932767 W02632 BE396786 R37261 AW207206 AW341473 AA448195 AI951341 AA249027 AL038984 AK001993 AL080066 AV652725 BE566226 AA345557 AA315222 AA090585 AA375688 AA301092 AA298454 W05762 AW607939 H51658 D83880 N84323 BE296821 AW947007 D61461 AW079261 AA329482 AW901780 AI354442 AA772275 R31663 AI354441 AI767525 H92431 AI916735 H93575 AI394255 AW014741 AI573090 C06195 AW612857 AW265195 AI339558 AI377532 AI308821 AI919424 AI589705 AW055215 AI336532 AI338051 AA806547 C75509 C00618 AW071172 AW769904 AA630381 AI678018 AI863985 D79662 BE221049 AW265018 AI589700 AW196655 N76573 AI370908 BE042393 N75017 AI698870 AW960115 AL133561 AL041090 AL117481 AL122069 AW439292 AI968826 AW072916 AI184913 AA489195 AW466994 AW469044 N59350 AI819642 AI280239 AI220572 AA789302 AI473611 AW841126 D60937 BE041395 AA491826 AA621946 AA715980 AA666102 AW970622 AA503009 AA502998 AA502989 AA502805 T92188 AA221036 R87170 BE537068 BE544757 C18935 AW812058 T92565 AA227415 AA233942 AA223237 AA668403 AA601627 AW869639 BE061833 BE000620 AW961170 AW847519 AA308542 AW821833 AW945688 C04699 AA205504 AA377241 AW821667 AA055720 AW817981 AW856468 AA155719 AA179928 T03007 AW754298 AA227407 AA113928 AA307904 C16859 AI798376 S46400 AW811617 AW811616 W00557 BE142245 AW858232 AW861851 AW858362 AA232351 AA218567 AA055556 AW858231 AW857541 AW814172 H66214 AW814398 AF134164 AA243093 AA173345 AA199942 AA223384 AA227092 AA227080 T12379 AA092174 T61139 AA149776 AA699829 AW879188 AW813567 AW813538 AI267168 AA157718 AA157719 AA100472 AA100774 AA130756 AA157705 AA157730 AA157715 AA053524 AW849581 AW854566 C05254 AW882836 T92637 AW812621 AA206583 AA209204 BE156909 AA226824 AI829309 AW991957 N66951 AA527374 H66215 AA045564 AI694265 H60808 AA149726 AW195620 BE081333 BE073424 AW817662 AW817705 AW817703 AW817659 BE081531 H59570 AA628980 AI126603 BE504035
5	412811	132943_1	
	413690	1383256_1	
	414883	15024_1	
10			
15			
20	415989	156454_1	
	417324	166714_1	
	418574	17690_1	
25	418712	1784125_1	
	419443	184788_1	
	419502	18535_1	
30			
35			
40			
45			
50	419936	189181_1	
	421582	2041_1	
55	422128	211994_1	
	423034	224122_1	
	423816	23234_1	
	424200	236595_1	
	424999	245835_1	
	426966	273896_1	
	426991	27415_1	
60	427260	276598_1	
	428023	28589_2	
65	429220	301384_1	
	429978	31150_1	
70			
75	430439	31808_1	
	430935	325772_1	
	431089	327825_1	
	431322	331543_1	
	432407	34624_1	
80	434414	38585_1	
85	436608	42361_3	



## WO 02/086443

## PCT/US02/12476

5	438091	44964_1	AW373062 T55662 AI299190 BE174210 AW579001 H01811 W40186 R67100 AI923886 AW952164 AA628440 AW898607 AW898616 AA709126 AW898628 AW898544 AA947932 AW898625 AW898622 AI276125 AI185720 AW510698 AA987230 T52522 BE467708 AW243400 AW043642 AI288245 AI186932 D52654 D55017 D52715 D52477 D53933 D54679 AI298739 AI146984 AI922204 N98343 BE174213 AA845571 AI813854 AI214518 AI635262 AI139455 AI707807 AI698085 AW884528 AI024768 AI004723 AW087420 AI565133 N94964 AI268939 AW513280 AI061126 AI435818 AI859106 AI360506 AI024767 AA513019 AA757598 X56196 AA902959 AI334784 AI860794 AA010207 AW890091 AW513771 AI951391 AI337671 T52499 AA890205 AI640908 H75966 AA463487 AA358688 AI961767 AI866295 AA780994 AI985913 BE174196 AA029094 AW592159 T55581 N79072 AI611201 AA910812 AI220713 AW149306 AI758412 AA045713 R79750 N76096 AW979121 AA847986 AA829098	
	439000	467716_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882 AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077	
	439285	47065_1	AL109688 R23665 R26578	
	439780	47673_1	AA570256 AW014761 AA573721 AI473237 AI022165 AA554071 AA127551 N90525 AW973623 AA447991 AA243852 BE328850 AI148171 AI359627 AI005068 AI356567 AA232991 AW016855 AA906902 AA233101 AA127550 BE512923	
	441128	51021_2	AI188710 AI032142 AW078833 N30308 AW675632 AI219028 AI341201 N22181 H95390 W24187 W24194 R17789	
10	443068	558874_1	Y10043 NM_005342 L05085 AL034450 BE614226 AW749053 AA379173 AA248230 BE514634 AA334622 R70656 AA367593 AA214649 AA369318 AW957081 R05760 AA039903 AI886597 AW630122 AA906264 AA041527 R01145 AI088688 BE463637 AA398795 AI354883 AI768938 AI569996 AI452952 AI168582 AI189869 AI086670 AW262560 AW613854 AA862839 AA435840 AA670197 AI024032 AI990659 AI990089 N81095 AA847919 AW960150 AA211075 AA044704 AA367594 AW582587 AW858854 AW818630 AW818281 AW818433 AW582595 AA096002 N83992	
	443947	586160_1	AI471630 BE540637 BE265481 AW407710 BE513882 BE546739 AA053597 BE140503 BE218514 AW956702 AI656234 AI636283 AI567265 AW340858 BE207794 AA053085 R69173 AA292343 AA454908 AA293504 AI659741 AI927478 AA399460 AI760441 AA346416 BE047245 AA730380 AA394063 AA454833 AI982791 AI567270 AI813332 AI767858 AA427705 D20284 AI221458 BE048537 AI263048 AA346417 AA911497 BE537702	
	447636	7301_1	AI638293 AW813561 AI761324 AW880941 AW880937	
	15	448993	79225_1	AW118072 AI631982 T15734 AA224195 AI701458 W20198 F26326 AA890570 N90552 AW071907 AI671352 AI375892 T03517 R88265 AI124088 AA224388 AI084316 AI354686 T33652 AI140719 AI720211 T03490 AI372637 T15415 AW205836 AA630384 T03515 T33230 AA017131 AA443303 T33623 AI222556 T33511 T33785 AI419606 D55612 W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW450652 AW449519 AA993634 AI806539 AA351618 AW449522 AI827626 AA904788 AA380381 AA886045 AA774409 BE003229 Z41756 AL133619 AA468118 AA383064 AI476447 T09430 AI673758 AA524895 AI581345 AI300820 AW498812 AA256162 AI559724 AI685732 AA602400 AA905453 AI204595 AW166541 AA157456 AA156269 AA383652 AA431072 AW592707 AI435410 AW272464 AI215594 AA622747 R74039 N35031 AI804128 AW513621 AA868351 AI026826 AI493388 AA614641 W81604 AI567080 AI214351 AA730140 AI125754 AI200813 AI269603 AI565082 AI807095 AI476629 AA505909 AI368449 AI686077 AI582930 AW085038 AA757863 AA730154 AI767072 AA468316 AI734130 AI734138 AA426284 AA433997 AI741241 AW043563 AI732741 AI732734 AA437369 AA425820 AA664048 R74130 BE144666 BE184942 AW238414 BE184946 AW993247 AW861464 AA203682 R11958
		449305	804424_1	BE550224 AA832519 N45402 AW885857 N29245 BE465409 W07677 AW970089 AI299731 AA482971 BE503548 H18151 W79223 AF086393 AA461301 W74510 R34182 AI090689 N46003 BE071550 R28075 AW134982 AI240204 AI138906 AW026179 AI572316 BE466182 AI206395 AI276154 AI273269 AI422817 AI371014 AI421274 AI188525 AA939164 BE549810 AW137865 AI694996 BE503841 AA459718 BE327407 BE467534 BE218421 BE467767 AA989054 BE467063 AI797130 BE327781
451105		859083_1		
451320		86576_1		
20		451807	8865_1	
	452410	9163_1		
	25	454241	1067807_1	
		455175	1257335_1	
		456237	168730_1	
458098		47395_1		
30				
35				
40				

TABLE 9C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
Nt\_position: Indicates nucleotide positions of predicted exons.

400512	9796593	Minus	1439-1615
400517	9796686	Minus	49996-50346
400560	9843598	Plus	94182-94323,97056-97243,101095-101236,102824-103005
400664	8118496	Plus	13558-13721,13942-14090,14554-14679
400665	8118496	Plus	16879-17023
400666	8118496	Plus	17982-18115,20297-20456
400749	7331445	Minus	9162-9293
400763	8131616	Minus	35537-35784
401027	7230983	Minus	70407-70554,71060-71160
401093	8516137	Minus	22335-23166
401203	9743387	Minus	172961-173056,173868-173928
401212	9858408	Plus	87839-88028
401411	7799787	Minus	144144-144329
401435	8217934	Minus	54508-55233
401464	6682291	Minus	170688-170834
401714	6715702	Plus	96484-96681
401747	9789672	Minus	118596-118816,119119-119244,119609-119761,120422-120990,130161-130381,130468-130593,131097-131258,131866-131932,132451-132575,133580-134011
401760	9929699	Plus	83126-83250,85320-85540,94719-95287
401780	7249190	Minus	28397-28617,28920-29045,29135-29296,29411-29567,29705-29787,30224-30573
401781	7249190	Minus	83215-83435,83531-83656,83740-83901,84237-84393,84955-85037,86290-86814
401785	7249190	Minus	165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
401797	6730720	Plus	6973-7118
401961	4581193	Minus	124054-124209
401985	2580474	Plus	61542-61750
401994	4153858	Minus	42904-43124,43211-43336,44607-44763,45199-45281,46337-46732
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
402260	3399665	Minus	113765-113910,115653-115765,116808-116940
402265	3287673	Plus	21059-21168
402297	6598824	Plus	35279-35405,35573-35659
402408	9796239	Minus	110326-110491



	402420	9796339	Plus	129750-129919
	402674	8077108	Minus	39290-39502
	402802	3287156	Minus	53242-53432
5	402994	2996643	Minus	4727-4969
	403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
	403306	8099945	Plus	127100-127251
	403329	8516120	Plus	96450-96598
	403381	9438267	Minus	26009-26178
10	403478	9958258	Plus	116458-116564
	403485	9966528	Plus	2888-3001,3198-3532,3655-4117
	403627	8569879	Minus	23868-24342
	403715	7239669	Plus	85128-85292
	404044	9558573	Minus	225757-225939
15	404076	9931752	Minus	3848-3967
	404101	8076925	Minus	125742-125997
	404140	9843520	Plus	37761-38147
	404165	9926489	Minus	69025-69128
	404185	4572584	Minus	129171-129327
20	404210	5006246	Plus	169926-170121
	404253	9367202	Minus	55675-56055
	404287	2326514	Plus	53134-53281
	404298	9944263	Minus	73591-73723
	404347	9838195	Plus	74493-74829
25	404440	7528051	Plus	80430-81581
	404721	9856648	Minus	173763-174294
	404794	4826439	Plus	101619-101898
	404854	7143420	Plus	14260-14537
	404877	1519284	Plus	1095-2107
30	404927	7342002	Plus	68690-69563
	404996	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40674,42351-42450
	405449	7622497	Plus	42236-42570
	405568	6006906	Plus	35912-36065
	405572	3800891	Plus	85230-85938
35	405646	4914350	Plus	741-969
	405676	4557087	Plus	73195-73917
	405770	2735037	Plus	61057-62075
	405932	7767812	Minus	123525-123713
	406137	9166422	Minus	30487-31058
40	406360	9256107	Minus	7513-7673
	406399	9256288	Minus	63448-63554
	406467	9795551	Plus	182212-182958

45 TABLE 10A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer and Non-malignant Lung Disease  
Table 2A shows about 307 genes up-regulated in non-malignant lung disease relative to lung tumors and normal body tissues and/or down-regulated in lung tumors relative to normal lung and non-malignant lung disease. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

50 Table 10B show the accession numbers for those Pkey's lacking UnigenelD's for table 10A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

55 Table 10C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 10A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

60 Pkey: Unique Eos probeset identifier number  
ExAccn: Exemplar Accession number, Genbank accession number  
UnigenelD: Unigene number  
Unigene Title: Unigene gene title  
R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

	Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2
65	404394			ENSP00000241075:TRRAP PROTEIN.	0.79	3.10
	404916			Target Exon	1.00	159.00
	405257			Target Exon	1.00	422.00
70	407228	M25079	Hs.155376	hemoglobin, beta	0.47	2.33
	407568	AA740964	Hs.62699	ESTs	1.00	123.00
	408562	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
	409031	AA376836	Hs.76728	ESTs	1.00	128.00
	410434	AF051152	Hs.63668	toll-like receptor 2	39.65	149.00
75	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	1.00	109.00
	410808	T40326	Hs.167793	ESTs	1.14	13.14
	412351	AL135960	Hs.73828	T-cell acute lymphocytic leukemia 1	0.37	2.27
	412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
	413795	AL040178	Hs.142003	ESTs	0.10	11.90
80	414154	AW205314	Hs.323060	ESTs	0.62	2.09
	414214	D49958	Hs.75819	glycoprotein M6A	0.03	4.55
	414998	NM_002543	Hs.77729	oxidised low density lipoprotein (lectin	0.64	2.97
	415122	D60708	Hs.22245	ESTs	0.07	8.97
	415765	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and	0.67	1.65
85	415775	H00747	Hs.29792	ESTs, Weakly similar to I38022 hypotheti	0.29	2.64
	415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00



5	416319	AI815601	Hs.79197	CD83 antigen (activated B lymphocytes, i	15.32	237.00
	416402	NM_000715	Hs.1012	complement component 4-binding protein,	0.64	4.00
	417355	D13168	Hs.82002	endothelin receptor type B	0.01	3.90
	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	36.30	357.00
	417511	AL049176	Hs.82223	chordin-like	1.00	179.00
10	418489	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h	0.02	6.00
	418726	BE241812	Hs.87860	protein tyrosine phosphatase, non-recept	1.00	113.00
	418741	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	0.44	1.90
	418883	BE387036	Hs.1211	acid phosphatase 5, tartrate resistant	0.96	2.04
	419086	NM_000216	Hs.89591	Kallmann syndrome 1 sequence	0.62	2.74
15	419150	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous	0.03	6.90
	419235	AW470411	Hs.288433	neurotrimin	1.48	5.13
	419407	AW410377	Hs.41502	hypothetical protein FLJ21276	37.55	336.00
	420556	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	0.80	3.65
	420656	AA279098	Hs.187636	ESTs	1.65	8.07
20	420729	AW964897	Hs.290825	ESTs	2.99	25.82
	421177	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	0.46	1.95
	422060	R20893	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.00	156.00
	422426	W79117	Hs.58559	ESTs	0.03	7.44
	422652	AW967969	Hs.118958	syntaxin 11	0.14	3.62
25	423099	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t	0.01	3.16
	424433	H04607	Hs.9218	ESTs	0.75	141.75
	424585	AA464840	Hs.131987	ESTs	1.00	167.00
	424711	NM_005795	Hs.152175	calcitonin receptor-like	0.43	3.01
	424973	X92521	Hs.154057	matrix metalloproteinase 19	0.37	19.45
30	425023	AW956889	Hs.154210	endothelial differentiation, sphingolipi	0.14	3.35
	425664	AJ006276	Hs.159003	transient receptor potential channel 6	1.00	94.00
	425998	AU076629	Hs.165950	fibroblast growth factor receptor 4	0.68	1.42
	426657	NM_015865	Hs.171731	solute carrier family 14 (urea transport	0.03	3.74
	426753	T89832	Hs.170278	ESTs	1.00	141.00
35	427558	D49493	Hs.2171	growth differentiation factor 10	1.00	117.00
	427983	M17706	Hs.2233	colony stimulating factor 3 (granulocyte	0.75	2.20
	428467	AK002121	Hs.184465	hypothetical protein FLJ11259	0.76	2.25
	428927	AA441837	Hs.90250	ESTs	0.01	3.62
	429496	AA453800	Hs.192793	ESTs	1.00	138.00
40	430468	NM_004673	Hs.241519	angiopoietin-like 1	1.00	132.00
	431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
	431728	NM_007351	Hs.268107	multimerin	1.00	157.00
	431848	AI378857	Hs.126758	ESTs, Highly similar to AF175283 1 zinc	0.34	2.24
	432128	AA127221	Hs.117037	ESTs	0.00	1.15
45	432519	AI221311	Hs.130704	ESTs, Weakly similar to BCHUIA S-100 pro	0.01	2.06
	433043	W57554	Hs.125019	lymphoid nuclear protein (LAF-4) mRNA	1.00	267.00
	433803	AI823593	Hs.27688	ESTs	1.00	105.00
	434730	AA644669	Hs.193042	ESTs	1.05	3.15
	435472	AW972330	Hs.283022	triggering receptor expressed on myeloid	0.83	1.94
50	436532	AA721522		gb:nv54h12.r1 NCI_CGAP_Ew1 Homo sapiens	1.00	218.00
	437119	AI379921	Hs.177043	ESTs	1.00	133.00
	437140	AA312799	Hs.283689	activator of CREM in testis	0.67	122.67
	437211	AA382207	Hs.5509	ecotropic viral integration site 2B	1.00	142.00
	437960	AI669586	Hs.222194	ESTs	1.00	147.00
55	438202	AW169287	Hs.22588	ESTs	1.00	141.00
	438873	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	0.71	3.66
	438875	AA827640	Hs.189059	ESTs	23.32	370.00
	441048	AA913488	Hs.192102	ESTs	0.77	8.50
	441188	AW292830	Hs.255609	ESTs	3.43	16.36
60	441499	AW298235	Hs.101689	ESTs	1.00	167.00
	444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
	444527	NM_005408	Hs.11383	small inducible cytokine subfamily A (Cy	46.47	153.00
	444561	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	0.01	3.08
	445279	R41900	Hs.22245	ESTs	0.60	141.00
65	446017	N98238	Hs.55185	ESTs	0.18	2.39
	446984	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15	0.10	2.16
	446998	N99013	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.01	2.53
	447357	AI375922	Hs.159367	ESTs	0.46	2.64
	448106	AI800470	Hs.171941	ESTs	18.05	296.00
70	448253	H25899	Hs.201591	ESTs	1.00	141.00
	449275	AW450848	Hs.205457	periastin	0.56	1.38
	450400	AI694722	Hs.279744	ESTs	0.88	4.33
	450696	AI654223	Hs.16026	hypothetical protein FLJ23191	0.52	2.08
	450726	AW204600	Hs.250505	retinoic acid receptor, alpha	0.79	2.01
75	451497	H83294	Hs.284122	Wnt inhibitory factor-1	0.35	2.03
	451533	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	0.13	2.25
	453636	R67837	Hs.169872	ESTs	1.00	116.00
	458332	AI000341	Hs.220491	ESTs	1.00	192.00
	459580	AA022888	Hs.176065	ESTs	0.20	2.98
80	400269			Eos Control	0.40	2.40
	403421			NM_016369*:Homo sapiens claudin 18 (CLDN	0.53	1.77
	407570	Z19002	Hs.37096	zinc finger protein 145 (Krueppel-like, e	0.01	3.18
	412295	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1	0.56	1.74
	414517	M24461	Hs.76305	surfactant, pulmonary-associated protein	0.64	1.50
85	417204	N81037	Hs.1074	surfactant, pulmonary-associated protein	0.33	1.16
	418307	U70867	Hs.83974	solute carrier family 21 (prostaglandin	0.53	1.55
	418935	T28499	Hs.89485	carbonic anhydrase IV	0.20	1.28
	421502	AF111856	Hs.105039	solute carrier family 34 (sodium phospho	0.78	1.90
	421798	N74880	Hs.29877	N-acylsphingosine amidohydrolase (acid c	0.59	1.54



5	423354	AB011130	Hs.127436	calcium channel, voltage-dependent, alph	0.59	1.55
	423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	0.35	1.62
	425438	T62216	Hs.270840	ESTs	0.23	9.45
	426828	NM_000020	Hs.172670	activin A receptor type II-like 1	0.03	1.71
10	427019	AA001732	Hs.173233	hypothetical protein FLJ10970	0.01	1.49
	428043	T92248	Hs.2240	uteroglobin	0.42	1.26
	430280	AA361258	Hs.237868	interleukin 7 receptor	0.46	2.43
	431433	X65018	Hs.253495	surfactant, pulmonary-associated protein	0.57	1.59
	431723	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.29	1.80
15	432985	T92363	Hs.178703	ESTs	0.32	2.27
	441835	AB036432	Hs.184	advanced glycosylation end product-speci	0.31	1.51
	442275	AW449467	Hs.54795	ESTs	0.55	1.78
	443709	AI082692	Hs.134662	ESTs	0.00	3.02
	444325	AW152618	Hs.16757	ESTs	0.32	2.49
20	450954	AI904740	Hs.25691	receptor (calcitonin) activity modifying	0.46	1.74
	451558	NM_001089	Hs.26630	ATP-binding cassette, sub-family A (ABC1	0.52	1.87
	453310	X70697	Hs.553	solute carrier family 6 (neurotransmitte	0.00	3.30
	456855	AF035528	Hs.153863	MAD (mothers against decapentaplegic, Dr	0.01	2.31
	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane	0.66	2.20
25	400754			Target Exon	1.00	297.00
	401045			C11001883*:gij6753278 ref NP_033938.1  c	1.00	109.00
	401083			NM_016582*:Homo sapiens peptide transpor	0.89	1.39
	402474			NM_004079:Homo sapiens cathepsin S (CTSS	1.45	4.47
	402808			ENSP00000235229:SEMB.	1.00	1.87
30	403021			C21000030:gij9955960 ref NP_063957.1  AT	1.00	149.00
	403438			NM_031419*:Homo sapiens molecule possess	1.06	2.96
	403687			NM_007037*:Homo sapiens a disintegrin-li	0.04	4.89
	403764			NM_005463:Homo sapiens heterogeneous nuc	1.00	225.00
	404277			NM_019111*:Homo sapiens major histocompa	0.97	1.93
35	404288			NM_002944*:Homo sapiens v-ros avian UR2	1.00	68.00
	404518	AI815601		CD83 antigen (activated B lymphocytes, i	0.02	1.83
	405106			C11001637*:gij5032241 ref NP_005732.1  z	1.00	235.00
	405381			Target Exon	1.00	93.00
	406387			Target Exon	1.37	6.02
40	406646	M33600		major histocompatibility complex, class	0.86	2.46
	406714	AI219304	Hs.266959	hemoglobin, gamma G	0.01	3.19
	406753	AA505665	Hs.217493	annexin A2	1.00	147.00
	406973	M34996	Hs.198253	major histocompatibility complex, class	1.03	2.04
	407248	U82275	Hs.94498	leukocyte immunoglobulin-like receptor,	1.00	64.00
45	407510	U96191		gb:Human trophoblast hypoxia-regulated f	1.00	90.00
	407731	NM_000066	Hs.38069	complement component 8, beta polypeptide	1.00	67.00
	407830	NM_001086	Hs.587	arylacetamide deacetylase (esterase)	1.00	102.00
	408045	AW138959	Hs.245123	ESTs	1.00	70.00
	408074	R20723		ESTs	1.00	112.00
50	408374	AW025430	Hs.155591	forkhead box F1	0.07	10.17
	409064	AA062954	Hs.141883	ESTs	0.39	2.31
	409083	AF050083	Hs.673	interleukin 12A (natural killer cell sti	1.00	95.00
	409153	W03754	Hs.50813	hypothetical protein FLJ20022	0.01	4.55
	409203	AA780473	Hs.687	cytochrome P450, subfamily IVB, polypept	0.01	3.72
55	409238	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp564G112 (fr	1.00	79.00
	409389	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	0.14	27.35
	409718	D86640	Hs.56045	src homology three (SH3) and cysteine ri	1.00	113.00
	410798	BE178622	Hs.16291	gb:PM3-HT0605-270200-001-a02 HT0605 Homo	0.64	2.47
	411020	NM_006770	Hs.67726	macrophage receptor with collagenous str	0.55	2.40
60	411667	BE160198		gb:QV1-HT0413-010200-059-h03 HT0413 Homo	1.00	111.00
	412000	AW576555	Hs.15780	ATP-binding cassette, sub-family A (ABC1	1.00	95.00
	412358	BE047490	Hs.24172	ESTs	1.00	87.00
	412420	AL035668	Hs.73853	bone morphogenetic protein 2	1.43	8.07
	412564	X83703	Hs.31432	cardiac ankyrin repeat protein	0.02	3.07
65	412869	AA290712	Hs.82407	CXC chemokine ligand 16	0.93	1.72
	412870	N22788	Hs.82407	CXC chemokine ligand 16	0.97	1.51
	413529	U11874	Hs.846	interleukin 8 receptor, beta	0.02	2.42
	413533	BE146973		gb:QV4-HT0222-011199-019-e05 HT0222 Homo	0.65	1.50
	413689	BE157286	Hs.20631	zinc finger protein, subfamily 1A, 5 (Pe	20.87	232.00
70	413724	AA131466	Hs.23767	hypothetical protein FLJ12666	1.00	80.00
	413800	AI129238	Hs.192235	ESTs	1.00	85.00
	413802	AW964490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-	1.00	213.00
	413829	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)	0.02	3.93
	414376	BE393856	Hs.66915	ESTs, Weakly similar to 16.7Kd protein [	1.00	115.00
75	414577	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to	0.49	1.94
	414700	H63202	Hs.38163	ESTs	0.03	3.75
	415078	AA311223	Hs.283091	found in inflammatory zone 3	0.86	1.95
	415120	N64464	Hs.34950	ESTs	1.00	120.00
	415323	BE269352	Hs.949	neutrophil cytosolic factor 2 (65kD, chr	0.60	2.48
80	415335	AA847758	Hs.111030	ESTs	1.00	95.00
	415582	W92445	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT	1.00	136.00
	416030	H15261	Hs.21948	ESTs	0.02	8.07
	416427	BE244050	Hs.79307	Rac/Cdc42 guanine exchange factor (GEF)	1.00	73.00
	416464	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	0.70	3.36
85	416585	X54162	Hs.79386	leiomodrin 1 (smooth muscle)	0.06	6.56
	416847	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.70	3.66
	417148	AA359896	Hs.293885	hypothetical protein FLJ14902	1.00	114.00
	417370	T28651	Hs.82030	tryptophanyl-tRNA synthetase	0.85	1.30
	417673	T87281	Hs.16355	ESTs	0.15	15.54



	418067	AI127958	Hs.83393	cystatin E/M	0.81	1.74
	418296	C01566	Hs.86671	ESTs	1.00	99.00
	418643	J03798	Hs.86948	small nuclear ribonucleoprotein D1 polyp	1.00	60.00
5	418832	X04011	Hs.88974	cytochrome b-245, beta polypeptide (chro	2.40	14.74
	418945	BE246762	Hs.89499	arachidonate 5-lipoxygenase	0.67	3.16
	419261	X07876	Hs.89791	wingless-type MMTV integration site fami	1.00	73.00
	419564	U08989	Hs.91139	solute carrier family 1 (neuronal/epithe	1.00	192.00
	419574	AK001989	Hs.91165	hypothetical protein	1.00	94.00
10	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	61.16	500.00
	420256	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula	0.52	1.70
	420285	AA258124	Hs.293878	ESTs, Moderately similar to ZN91_HUMAN Z	1.00	172.00
	420577	AA278436	Hs.186649	ESTs	1.00	97.00
	421262	AA286746	Hs.9343	Homo sapiens cDNA FLJ14265 fis, clone PL	1.00	64.00
15	421445	AA913059	Hs.104433	Homo sapiens, clone IMAGE:4054868, mRNA	0.88	1.51
	421470	R27496	Hs.1378	annexin A3	0.05	11.26
	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	1.00	73.00
	421563	NM_006433	Hs.105806	granulysin	0.82	2.42
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	5.50	31.57
20	421855	F06504	Hs.27384	ESTs, Moderately similar to ALU4_HUMAN A	1.00	129.00
	421913	AI934365	Hs.109439	osteoglycin (osteoinductive factor, mime	1.00	101.00
	421952	AA300900	Hs.98849	ESTs, Moderately similar to AF161511 1 H	0.60	63.60
	422232	D43945	Hs.113274	transcription factor EC	1.00	148.00
	422386	AF105374	Hs.115830	heparan sulfate (glucosamine) 3-O-sulfo	1.40	3.98
25	423168	R34385	Hs.124940	GTP-binding protein	0.34	3.59
	423196	AK001866	Hs.125139	hypothetical protein FLJ11004	0.55	2.00
	423387	AJ012074		vasoactive intestinal peptide receptor 1	0.09	2.13
	423424	AF150241	Hs.128433	prostaglandin D2 synthase, hematopoietic	1.00	141.00
	423456	AL110151	Hs.128797	DKFZP586D0824 protein	1.00	66.00
30	423696	Z92546		Sushi domain (SCR repeat) containing	0.73	1.27
	424027	AW337575	Hs.201591	ESTs	0.54	2.58
	424212	NM_005814	Hs.143131	glycoprotein A33 (transmembrane)	0.77	2.47
	425087	R62424	Hs.126059	ESTs	1.00	74.00
	425175	AF020202	Hs.155001	UNC13 (C. elegans)-like	0.85	1.96
35	425771	BE561776	Hs.159494	Bruton agammaglobulinemia tyrosine kinas	1.18	2.56
	426486	BE178285	Hs.170056	Homo sapiens mRNA; cDNA DKFZp586B0220 (f	1.00	76.00
	427507	AF240467	Hs.179152	toll-like receptor 7	1.00	63.00
	427618	NM_000760	Hs.2175	colony stimulating factor 3 receptor (gr	0.60	2.19
	427732	NM_002980	Hs.2199	secretin receptor	0.97	1.42
40	427952	AA765368	Hs.293941	ESTs, Moderately similar to A53959 throm	1.00	105.00
	428709	BE268717	Hs.104916	hypothetical protein FLJ21940	1.00	80.00
	428769	AW207175	Hs.106771	ESTs	0.09	2.55
	428780	AI478578	Hs.50636	ESTs	1.00	98.00
	428833	AI928355	Hs.185805	ESTs	1.00	113.00
45	429657	D13626	Hs.2465	KIAA0001 gene product; putative G-protei	1.00	52.00
	430212	AA469153		gb:nc67f04.s1 NCI_CGAP_Pr1 Homo sapiens	1.00	132.00
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	0.11	15.60
	430376	AW292053	Hs.12532	chromosome 1 open reading frame 21	1.00	103.00
	430414	AW365665	Hs.120388	ESTs	0.50	6.96
50	430656	AA482900	Hs.162080	ESTs	1.00	70.00
	430843	AI734149	Hs.119514	ESTs	1.00	90.00
	430998	AF128847	Hs.204038	indolethylamine N-methyltransferase	0.29	1.84
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	1.00	79.00
	431921	N46466	Hs.58879	ESTs	0.91	1.67
55	432176	AW090386	Hs.112278	arrestin, beta 1	0.66	2.63
	432203	AA305746	Hs.49	macrophage scavenger receptor 1	1.00	76.00
	432231	AA339977	Hs.274127	CLST 11240 protein	0.46	1.46
	432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	0.79	2.25
	432522	D11466	Hs.51	phosphatidylinositol glycan, class A (pa	1.93	4.83
60	432596	AJ224741	Hs.278461	matrilin 3	0.04	5.79
	432850	X87723	Hs.3110	angiotensin receptor 2	1.00	167.00
	433138	AB029496	Hs.59729	semaphorin sem2	0.04	9.16
	433563	AI732637	Hs.277901	ESTs	1.00	91.00
	433588	AI056872	Hs.133386	ESTs	120.16	315.00
65	434445	AI349306	Hs.11782	ESTs	0.60	1.84
	435496	AW840171	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	128.00
	435974	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor	1.00	108.00
	436061	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	1.00	91.00
	437157	BE048860	Hs.120655	ESTs	1.00	87.00
70	437207	T27503	Hs.15929	hypothetical protein FLJ12910	1.00	105.00
	437311	AA370041	Hs.9456	SWI/SNF related, matrix associated, acti	1.00	71.00
	437439	H29796	Hs.269622	ESTs	1.00	115.00
	438199	AW016531	Hs.122147	ESTs	1.00	80.00
	439551	W72062	Hs.11112	ESTs	0.30	3.10
75	440515	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene famil	1.00	77.00
	440887	AI799488	Hs.135905	ESTs	1.00	85.00
	441025	AA913880	Hs.176379	ESTs	1.00	82.00
	441384	AA447849	Hs.288660	Homo sapiens cDNA: FLJ22182 fis, clone H	0.79	1.89
	441735	AI738675	Hs.127346	ESTs	1.00	75.00
80	442200	AW590572	Hs.235768	ESTs	0.78	5.83
	442832	AW206560	Hs.253569	ESTs	0.03	10.88
	442957	AI949952	Hs.49397	ESTs	1.00	70.00
	443282	T47764	Hs.132917	ESTs	1.00	197.00
	443547	AW271273	Hs.23767	hypothetical protein FLJ12666	1.00	253.00
85	443951	F13272	Hs.111334	ferritin, light polypeptide	0.55	2.09
	444330	AI597655	Hs.49265	ESTs	1.00	90.00



5	444515	AW204908	Hs.169979	ESTs	1.00	84.00
	445769	AI741471	Hs.23666	ESTs	0.02	4.38
	445908	R13580	Hs.13436	Homo sapiens clone 24425 mRNA sequence	1.00	97.00
	446291	BE397753	Hs.14623	interferon, gamma-inducible protein 30	0.93	1.69
	446917	AI347863	Hs.156672	ESTs	1.00	106.00
10	447261	NM_006691	Hs.17917	extracellular link domain-containing 1	0.40	47.20
	447432	AW958473	Hs.301957	nudix (nucleoside diphosphate linked moi	1.00	100.00
	447482	AB033059	Hs.18705	KIAA1233 protein	0.05	8.21
	447997	H00656	Hs.29792	ESTs, Weakly similar to I38022 hypotheti	0.02	5.42
	448299	AA497044	Hs.20887	hypothetical protein FLJ10392	1.00	79.00
15	448782	AL050295	Hs.22039	KIAA0758 protein	0.42	1.56
	450575	NM_005859	Hs.29117	purine-rich element binding protein A	0.17	11.33
	450584	AA040403	Hs.60371	ESTs	1.00	94.00
	450693	AW450461	Hs.203965	ESTs	1.00	91.00
	450715	AI266484	Hs.31570	ESTs, Weakly similar to KIAA1324 protein	1.00	152.00
20	451103	R52804	Hs.25956	DKFZP564D206 protein	1.00	86.00
	451220	AF124251	Hs.26054	novel SH2-containing protein 3	0.60	1.30
	451668	Z43948	Hs.326444	cartilage acidic protein 1	0.54	1.91
	452197	AW023595	Hs.232048	ESTs	1.00	67.00
	452331	AA598509	Hs.29117	purine-rich element binding protein A	4.53	11.07
25	452353	C18825	Hs.29191	epithelial membrane protein 2	0.72	2.24
	453049	BE537217	Hs.30343	ESTs	1.00	68.00
	453107	NM_016113	Hs.279746	vanilloid receptor-like protein 1	0.83	1.70
	453355	AW295374	Hs.31412	Homo sapiens cDNA FLJ11422 fis, clone HE	1.00	132.00
	453390	AA862496	Hs.28482	ESTs	1.00	72.00
30	453531	AA417940		ESTs, Weakly similar to JC5795 CDEP prot	1.00	68.00
	454741	BE154396		gb:CM2-HT0342-091299-050-b05 HT0342 Homo	0.57	2.89
	456579	AA287827	Hs.284205	up-regulated by BCG-CWS	1.00	82.00
	456672	AK002016	Hs.114727	Homo sapiens, clone MGC:16327, mRNA, com	0.79	1.96
	457400	AF032906	Hs.252549	cathepsin Z	1.03	3.25
35	457718	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S	1.00	113.00
	459696	F03027		gb:HSC1KA072 normalized infant brain cDN	1.00	544.00

TABLE 10B

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

40	Pkey	CAT Number	Accession
	408074	103684_1	R20723 AA263003 AA333976 AA334725 AA334151 AW965490 AA310513 AI810530 D31302 AW134897 AA830127 AA046953 AI668930 C06094 AW104534
45	411667	1253334_1	BE160198 AW935898 T11520 AW935930 AW856073 AW861034
	413533	1375344_1	BE146973 BE146972 BE147042 BE147018 BE146783 BE147020 BE146781 BE147019 BE146766 BE147021 BE146952 BE146767 BE147044 BE146797 BE146776 BE146985 BE146793 BE146768 BE146771 BE146954 BE146760 BE147048 BE147025 BE147030
50	423387	22779_1	AJ012074 U11087 L13288 X75299 L20295 AW630780 H14880 T28037 AI872991 R72136 AW449839 T81622 T79697 T29519 R94105 T83923 R73300 AI797007 R73390 AA961010 H74168 AI689932 BE045543 AI808418 AI608912 AI806573 AW884084 AW872978 AW872985 AA565655 AI022915 R50647 R73210 H45098 R46451 AW166269 T71132 AI264547 R52146 AI304920 R73391 AW884059 AW884085 H73241 T60038 T79612 R73145 R50549 AI094557 AI668793 R72302 AI564366 W01956 AA418962 W32571 R72840 H45409 R72085 R46356 R46758 AA508805 AA418798 T83751 R94072 T16182 AA928785 AA903896
	423696	23112_1	Z92546 AA330586 AI570568 AW341487 AI827050 AW298668 AI792189 AI015693 AI733599 AI572251 AI672488 AW193262 AI244716 AI864375 AI206100 AA912444 AI269365 AI640254 AW772466 AI867336 AA627604 H16914 AA358477 AA338009
55	430212	314437_1	AA469153 AI718503 AA469225
	436532	421802_1	AA721522 AW975443 T93070
60	453531	97026_1	AA417940 AA036735 T07025
	454741	1232559_1	BE154396 AW817959 BE154393

TABLE 10C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
Nt\_position: Indicates nucleotide positions of predicted exons.

70	Pkey	Ref	Strand	Nt_position
	400754	7331445	Plus	144559-144684
75	401045	8117619	Plus	90044-90184,91111-91345
	401083	3242744	Plus	33192-33360
80	402474	7547175	Minus	53526-53628,55755-55920,57530-57757
	402808	6456148	Minus	114964-115136,115461-115585,115931-116047,117666-117771,118004-118102
85	403021	7547270	Plus	120799-120966
	403421	9665041	Minus	126609-126773,139986-140205
90	403438	9719679	Plus	90792-90938
	403687	7387384	Plus	9009-9534
95	403764	7717105	Minus	118692-118853
	404277	1834458	Minus	91665-91946
100	404288	2769644	Plus	3512-3691
	404394	3135305	Minus	37121-37205,37491-37762,41053-41140,41322-41593,41773-41919
105	404518	8151988	Plus	84494-84603
	404916	7341826	Plus	91057-91188
110	405106	8079395	Minus	80877-81418
	405257	7329310	Plus	73121-73273
115	405381	6006920	Minus	7636-8054







TABLE 11A: Genes Distinguishing Adenocarcinoma from Other Lung Diseases and Normal Lung

Table 11A shows about 84 genes upregulated in lung adenocarcinomas relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 11B show the accession numbers for those Pkey's lacking UnigeneID's for table 11A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 11C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 11A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigeneID: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
403329			Target Exon	1.00	61.00
406399			NM_003122*:Homo sapiens serine protease	1.00	39.00
406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	226.37	350.00
407869	AI827976	Hs.24391	hypothetical protein FLJ13612	0.77	1.18
407881	AW072003	Hs.40968	heparan sulfate (glucosamine) 3-O-sulfot	1.00	10.00
408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
409187	AF154830	Hs.50966	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
410076	T05387	Hs.7991	ESTs	1.12	1.50
410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
410399	BE068889		synuclein, gamma (breast cancer-specific	0.92	1.06
411908	L27943	Hs.72924	cytidine deaminase	1.00	1.00
412612	NM_000047	Hs.74131	arylsulfatase E (chondrodysplasia puncta	1.02	1.03
414075	U11862	Hs.75741	amiloride binding protein 1 (amine oxida	0.84	1.07
416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
417542	J04129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
419502	AU076704		fibrinogen, A alpha polypeptide	13.05	115.00
419631	AW188117	Hs.303154	popeye protein 3	1.00	13.00
420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
421190	U95031	Hs.102482	mucin 5, subtype B, tracheobronchial	1.17	1.55
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase 1, l	1.00	3.00
421582	AI910275		trefoil factor 1 (breast cancer, estroge	1.23	1.00
422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
422095	AI868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
423472	AF041260	Hs.129057	breast carcinoma amplified sequence 1	48.13	72.00
423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
424502	AF242388	Hs.149585	lengsin	1.00	1.00
424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	1.00	59.00
424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
424960	BE245380	Hs.153952	5' nucleotidase (CD73)	1.00	1.00
425523	AB007948	Hs.158244	KIAA0479 protein	1.00	35.00
426230	AA367019	Hs.241395	protease, serine, 1 (trypsin 1)	1.00	83.00
427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
428585	AB007863	Hs.185140	KIAA0403 protein	1.00	6.00
428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	1.59	1.69
430508	AI015435	Hs.104637	ESTs	4.75	7.27
430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
431548	AI834273	Hs.9711	novel protein	5.66	15.00
431566	AF176012	Hs.260720	J domain containing protein 1	49.76	37.00
431986	AA536130	Hs.149018	Novel human gene mapping to chromosome 20	1.19	1.47
432375	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
432677	NM_004482	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
433819	AW511097	Hs.112765	ESTs	3.71	8.00
434001	AW950905	Hs.3697	serine (or cysteine) proteinase inhibito	29.31	72.00
434424	AI811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
434792	AA649253	Hs.132458	ESTs	8.52	44.00
436217	T53925	Hs.107	fibrinogen-like 1	57.97	31.00
436749	AA584890	Hs.5302	lectin, galactoside-binding, soluble, 4	1.10	1.41
436972	AA284679	Hs.25640	claudin 3	1.59	1.46
437866	AA156781		metallothionein 1E (functional)	3.62	101.00
437935	AW939591	Hs.5940	mucin 13, epithelial transmembrane	1.60	1.39
438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	1.00	1.00
439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00



5	439759	AL359055	Hs.67709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00
	441377	BE218239	Hs.202656	ESTs	22.03	1.00
	443614	AV655386	Hs.7645	fibrinogen, B beta polypeptide	1.00	16.00
	443813	AA876372	Hs.93961	Homo sapiens mRNA; cDNA DKFZp667D095 (fr	1.20	1.99
10	443991	NM_002250	Hs.10082	potassium intermediate/small conductance	5.71	6.87
	444670	H58373	Hs.332938	hypothetical protein MGC5370	1.98	38.00
	444931	AV652066	Hs.75113	general transcription factor IIIA	1.00	54.00
	446102	AW168067	Hs.317694	ESTs	1.00	1.00
	446163	AA026880	Hs.25252	Homo sapiens cDNA FLJ13603 fis, clone PL	1.00	36.00
15	446469	BE094848	Hs.15113	homogentisate 1,2-dioxygenase (homogenti	1.00	11.00
	447388	AW630534	Hs.76277	Homo sapiens, clone MGC:9381, mRNA, comp	1.24	1.16
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
	448243	AW369771	Hs.52620	integrin, beta 8	15.84	1.00
	448844	AI581519	Hs.177164	ESTs	1.00	31.00
20	449444	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	1.00	83.00
	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	452689	F33868	Hs.284176	transferrin	1.54	1.44
	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00
	453464	AI884911	Hs.32989	receptor (calcitonin) activity modifying	1.55	2.45
	453735	AI066629	Hs.125073	ESTs	1.01	1.30

TABLE 11B

25	Pkey: Unique Eos probeset identifier number		
	CAT number: Gene cluster number		
	Accession: Genbank accession numbers		
30	Pkey	CAT Number	Accession
	410399	11995_1	BE068889 BE068882 AF044311 AF017256 NM_003087 AF037207 AF010126 AA633976 AA872836 BE298825 BE299889 AI016464 AI684600 AI936527 AA804675 AA394097 AI139933 AA946606 BE171313 AA722407 AA293803 AI468480 AA056035 AA055968 AW796957 AI637713 AA410737 H49348 AA486472 AA411094 AA235594 AA402624 AA443638 AW452137 AA421708 AW265211 AI493266 AA365132 AW966044 AU076704 T74854 T74860 T72098 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72826 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M64982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 R06796 H48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AA011465 AA345378 AV654847 AV654272 AV656001 AI064740 T82897 N33594 AA344542 AW805054 AI207457 T61743 AA026737 H94389 AA382695 AA918409 T68044 S82092 T39959 AI017721 AA312395 AA312919 T40156 H66239 AV652989 H38728 R98521 AV655200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343596 AW470774 AV651256 N54417 AA812862 AW182929 AI111192 H61463 H72060 AA344503 H38639 AI277511 AV661108 AI207625 T47810 AA235252 T27853 T47778 R95746 H70620 AA701463 AW827166 R98475 C20925 AV657287 T71959 T71313 T73920 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T27854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 AI110639 AA344603 AF063513 T64696 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74677 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA342807 AA341964 T53747 T72042 T62764 AI064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 AI207557 T68219 W86031 T69081 T64232 R93196 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 AI248502 R29454 T64764 T57001 T73052 T71429 T51176 T58866 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 AI910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 AI571948 AA507595 AA614579 AA587613 R83818 AA568312 AA614409 AA307578 AI925552 AW950155 AI910083 M12075 BE074052 AW004668 AA578674 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA858276 R55389 AI001051 AW050700 AW750216 AA614539 BE074045 AI307407 AW602303 BE073575 AI202532 AA524242 AI970839 AI909751 BE076078 AI909749 R55292 AA156781 AW293839 U52054 AA024963 AA778446 BE073977 AW444904 AW602574 BE164040 BE164012 BE163972 BE163974 BE163992 AA837481 AW468444 BE185091 AW468002 AA687333 AA811830 AA581806 AI866686 AI572124 AA043777 AA040926 D20160 AI536733 AA812489 AW874142 AI471883 W84421 AA156850 W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW450652 AW449519 AA993634 AI806539 AA351618 AW449522 AI827626 AA904788 AA380381 AA886045 AA774409 BE003229 Z41756

TABLE 11C

70	Pkey: Unique number corresponding to an Eos probeset		
	Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.		
75	Strand: Indicates DNA strand from which exons were predicted.		
	Nt_position: Indicates nucleotide positions of predicted exons.		

Pkey	Ref	Strand	Nt_position
403329	8516120	Plus	96450-96598
406399	9256288	Minus	63448-63554



TABLE 12A: Genes Distinguishing Squamous Cell Carcinoma from Other Lung Diseases and Normal Lung

Table 12A shows about 72 genes upregulated in squamous cell carcinomas of the lung relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 12B show the accession numbers for those Pkey's lacking UnigenelD's for table 12A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 12C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 12A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400666			NM_002425:Homo sapiens matrix metallopro	3.26	3.22
401780			NM_005557*:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275*:Homo sapiens keratin 15 (KRT1	4.13	2.70
401994			Target Exon	61.84	47.00
402075			ENSP00000251056*:Plasma membrane calcium	1.00	1.00
404996			Target Exon	1.00	1.00
407839	AA045144	Hs.161566	ESTs	173.91	108.00
408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	151.17	8.00
408522	AI541214	Hs.46320	Small proline-rich protein SPRK [human,	1.98	1.24
410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
417366	BE185289	Hs.1076	small proline-rich protein 1B (comifin)	8.97	3.27
418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
420783	AI659838	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
421773	W69233	Hs.112457	ESTs	1.12	1.14
421948	L42583	Hs.334309	keratin 6A	51.83	20.25
421978	AJ243662	Hs.110196	NICE-1 protein	1.01	0.91
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	233.42	68.00
424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito	1.00	1.00
424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00
427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	11.80	1.00
430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
430890	X54232	Hs.2699	glypican 1	1.58	1.40
431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
431846	BE019924	Hs.271580	uroplakin 1B	4.49	2.51
433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
434360	AW015415	Hs.127780	ESTs	40.98	27.00
434880	U02388	Hs.101	cytochrome P450, subfamily IVF, polypept	1.00	1.00
435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
436511	AA721252	Hs.291502	ESTs	16.76	14.00
438403	AA806607	Hs.292206	ESTs	1.00	1.00
439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
439606	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
439706	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	86.55	11.00
440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
441525	AW241867	Hs.127728	ESTs	1.53	1.42
443162	T49951	Hs.9029	DKFZP434G032 protein	31.11	38.00
444378	R41339	Hs.12569	ESTs	1.00	1.00



446292	AF081497	Hs.279682	Rh type C glycoprotein	1.55	1.26
447078	AW885727	Hs.9914	ESTs	47.24	24.00
447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00
450832	AW970602	Hs.105421	ESTs	25.17	36.00
452240	AI591147	Hs.61232	ESTs	13.42	1.00
453317	NM_002277	Hs.41696	keratin, hair, acidic,1	1.19	1.27
453830	AA534296	Hs.20953	ESTs	24.92	25.00
454098	W27953	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11
455601	AI368680	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00

TABLE 12B

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT Number	Accession
439285	47065_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882 AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077

TABLE 12C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
Nt\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
400666	8118496	Plus	17982-18115,20297-20456
401780	7249190	Minus	28397-28617,28920-29045,29135-29296,29411-29567,29705-29787,30224-30573
401781	7249190	Minus	83215-83435,83531-83656,83740-83901,84237-84393,84955-85037,86290-86814
401785	7249190	Minus	165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
401994	4153858	Minus	42904-43124,43211-43336,44607-44763,45199-45281,46337-46732
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
404996	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40674,42351-42450



TABLE 13A: Genes Distinguishing Non-Malignant Lung Disease from Lung Tumors and Normal lung

Table 13A shows about 23 genes upregulated in non-malignant lung disease relative to lung tumors and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 13B show the accession numbers for those Pkey's lacking UnigenelD's for table 13A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 13C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 13A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
ExAccn: Exemplar Accession number, Genbank accession number  
UnigenelD: Unigene number  
Unigene Title: Unigene gene title  
R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2
408562	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
409031	AA376836	Hs.76728	ESTs	1.00	128.00
412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00
417511	AL049176	Hs.82223	chordin-like	1.00	179.00
418819	AA228776	Hs.191721	ESTs	1.00	140.00
422060	R20893	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.00	156.00
424585	AA464840	Hs.131987	ESTs	1.00	167.00
426753	T89832	Hs.170278	ESTs	1.00	141.00
429496	AA453800	Hs.192793	ESTs	1.00	138.00
430719	AA488988	Hs.293796	ESTs	1.00	133.00
431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
431728	NM_007351	Hs.268107	multimerin	1.00	157.00
436532	AA721522		gb:nv54h12.r1 NCI_CGAP_Ew1 Homo sapiens	1.00	218.00
437960	AI669586	Hs.222194	ESTs	1.00	147.00
438202	AW169287	Hs.22588	ESTs	1.00	141.00
441499	AW298235	Hs.101689	ESTs	1.00	167.00
444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
448253	H25899	Hs.201591	ESTs	1.00	141.00
453636	R67837	Hs.169872	ESTs	1.00	116.00
458332	AI000341	Hs.220491	ESTs	1.00	192.00
459587	AA031956		gb:zk15e04.s1 Soares_pregnant_uterus_NbH	1.00	154.00

TABLE 13B

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT Number	Accession
431089	327825_1	BE041395 AA491826 AA621946 AA715980 AA666102
436532	421802_1	AA721522 AW975443 T93070

TABLE 13C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
NL\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076



TABLE 14A: Preferred Utility and Subcellular Localization for Potential Lung Disease Targets

Table 14A shows the subcellular localization and preferred utility for the genes appearing in Tables 9A and 10A. mAb symbolizes monoclonal antibody, diag symbolizes diagnostic, s.m. symbolizes small molecule, and CTL symbolizes cytotoxic lymphocytic ligand. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 14B show the accession numbers for those Pkey's lacking UnigenelD's for table 14A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 14C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 14A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number

ExAccn: Exemplar Accession number, Genbank accession number

UnigenelD: Unigene number

Unigene Title: Unigene gene title

Pref. Utility: Preferred Utility

Pred. Loc: Predicted subcellular localization

Pkey	ExAccn	UnigenelD	Unigene Title	Pref Utility	Pred. Loc
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	mAb & diag & s.m.	extracellular
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	mAb	plasma membrane
402075			ENSP00000251056*:Plasma membrane calcium	mAb & diag	secreted
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	diag	secreted
408243	Y00787	Hs.624	interleukin 8	diag	secreted
408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	mAb & s.m.	plasma membrane
408908	BE296227	Hs.250822	serine/threonine kinase 15	s.m.	cytoplasm
409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	CTL & diag	secreted
409103	AF251237	Hs.112208	XAGE-1 protein	CTL	nuclear
409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	diag	secreted
409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	diag	secreted
409757	NM_001898	Hs.123114	cystatin SN	diag	extracellular
409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	CTL	nuclear
409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	diag	extracellular
410001	AB041036	Hs.57771	kallikrein 11	diag	extracellular
410407	X66839	Hs.63287	carbonic anhydrase IX	mAb & s.m.	plasma membrane
410418	D31382	Hs.63325	transmembrane protease, serine 4	mAb & diag & s.m.	plasma membrane
412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines	s.m.	
412719	AW016610	Hs.816	ESTs	s.m.	nuclear
414774	X02419	Hs.77274	plasminogen activator, urokinase	diag	extracellular
414883	AA926960		CDC28 protein kinase 1	s.m.	
415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	CTL & diag	extracellular
415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibito	mAb & diag & s.m.	secreted
415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	mAb & s.m.	plasma membrane
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	diag	extracellular
417034	NM_006183	Hs.80962	neurotensin	diag	extracellular
417079	U65590	Hs.81134	interleukin 1 receptor antagonist	diag	extracellular
417308	H60720	Hs.81892	KIAA0101 gene product	s.m.	mitochondrial
417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	mAb & diag	secreted
417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	mAb	plasma membrane
417933	X02308	Hs.82962	thymidylate synthetase	s.m.	endoplasmic reticulum
418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	s.m.	cytoplasm
418506	AA084248	Hs.85339	G protein-coupled receptor 39	mAb & s.m.	plasma membrane
418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)	CTL	cytoplasmic
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	diag	secreted
419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor t	mAb & s.m.	plasma membrane
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	CTL & s.m.	mitochondrial
419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy	diag	secreted
419235	AW470411	Hs.288433	neurotrimin	mAb & diag	plasma membrane
419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	mAb & s.m.	plasma membrane
419556	U29615	Hs.91093	chitinase 1 (chitotriosidase)	mAb & diag	extracellular*
420610	AI683183	Hs.99348	distal-less homeo box 5	CTL	nuclear
421110	AJ250717	Hs.1355	cathepsin E	sm & diag	extracellular
421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	diag	secreted
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	mAb & s.m.	plasma membrane
421552	AF026692	Hs.105700	secreted frizzled-related protein 4	diag	secreted
421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	mAb & s.m.	plasma membrane
421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	mAb & s.m.	plasma membrane
422109	S73265	Hs.1473	gastrin-releasing peptide	diag	secreted
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	diag	secreted
422282	AF019225	Hs.114309	apolipoprotein L	diag	secreted
422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	s.m.	nuclear
422424	AI186431	Hs.296638	prostate differentiation factor	diag	extracellular
422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	s.m.	cytoplasm
422809	AK001379	Hs.121028	hypothetical protein FLJ10549	s.m.	nuclear
422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	diag	extracellular
422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	CTL & s.m.	
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	diag	
423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	mAb & diag & s.m.	secreted
423961	D13666	Hs.136348	periostin (OSF-2os)	mAb & diag	extracellular
424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito	diag	secreted
424381	AA285249	Hs.146329	protein kinase Chk2	s.m.	nuclear



5	424502	AF242388	Hs.149585	lengsin	s.m.	cytoplasmic
	424503	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,	mAb & s.m.	plasma membrane
	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	diag	extracellular
	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	mAb & diag & s.m.	secreted
	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	s.m.	cytoplasmic
10	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	mAb	plasma membrane
	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	s.m.	
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	mAb & diag	plasma membrane
	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	mAb & s.m.	plasma membrane
	426215	AW963419	Hs.155223	stanniocalcin 2	mAb & diag	secreted
15	426427	M86699	Hs.169840	TTK protein kinase	CTL & s.m.	nuclear
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	mAb & diag	secreted
	427335	AA448542	Hs.251677	G antigen 7B	CTL	cytoplasmic
	427747	AW411425	Hs.180655	serine/threonine kinase 12	s.m.	cytoplasmic
	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	diag	
20	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	mAb & diag & s.m.	extracellular
	428450	NM_014791	Hs.184339	KIAA0175 gene product	s.m.	nuclear
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	s.m.	nuclear
	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	mAb & s.m.	plasma membrane
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	CTL & s.m.	nuclear
25	428698	AA852773	Hs.334838	KIAA1866 protein	mAb	
	428748	AW593206	Hs.98785	Ksp37 protein	diag	extracellular
	428758	AA433988	Hs.98502	CA125 antigen; mucin 16	diag	mitochondria*
	428969	AF120274	Hs.194689	artemin	diag	extracellular
	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	mAb & s.m.	plasma membrane
30	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	mAb & s.m.	plasma membrane
	429547	AW009166	Hs.99376	ESTs	diag	secreted
	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	mAb & diag	secreted
	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	s.m.	
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	mAb & s.m.	plasma membrane
35	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	diag	extracellular
	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	mAb & s.m.	plasma membrane
	431846	BE019924	Hs.271580	uroplakin 1B	mAb & diag	plasma membrane
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	mAb & diag	plasma membrane
	432201	AI538613	Hs.298241	Transmembrane protease, serine 3	mAb & diag & s.m.	plasma membrane
40	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	s.m.	nuclear
	435505	AF200492	Hs.211238	interleukin-1 homolog 1	diag	secreted
	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	s.m.	
	437016	AU076916	Hs.5398	guanine monophosphate synthetase	s.m.	cytoplasm
	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an	CTL	ER
45	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	CTL	nuclear
	437852	BE001836	Hs.256897	ESTs, Weakly similar to DJ365O12.1 [H.sa	mAb & s.m.	plasma membrane
	439223	AW238299	Hs.250618	UL16 binding protein 2	mAb	plasma membrane
	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	mAb & s.m.	
	439606	W79123	Hs.58561	G protein-coupled receptor 87	mAb & s.m.	plasma membrane
50	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	mAb & s.m.	plasma membrane
	440006	AK000517	Hs.6844	NALP2 protein; PYRIN-Containing APAF1-li	s.m.	nuclear
	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	s.m.	
	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447	mAb & s.m.	plasma membrane
	443247	BE614387	Hs.333893	c-Myc target JPO1	CTL	extracellular*
55	443426	AF098158	Hs.9329	chromosome 20 open reading frame 1	CTL	
	443859	NM_013409	Hs.9914	folliculin	diag	extracellular
	444006	BE395085	Hs.10086	type I transmembrane protein Fn14	mAb	plasma membrane
	444371	BE540274	Hs.239	forkhead box M1	s.m.	nuclear
	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti	diag	secreted
60	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	mAb & diag	plasma membrane
	445537	AJ245671	Hs.12844	EGF-like-domain, multiple 6	mAb & diag	secreted
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	diag	secreted
	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	diag	extracellular
	447033	AI357412	Hs.157601	ESTs	CTL & diag	secreted
65	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	CTL	
	448243	AW369771	Hs.52620	integrin, beta 8	mAb & s.m.	plasma membrane
	448844	AI581519	Hs.177164	ESTs	mAb & s.m.	
	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	mAb	plasma membrane
	449722	BE280074	Hs.23960	cyclin B1	s.m.	cytoplasm
70	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte	mAb & s.m.	plasma membrane
	450375	AA009647		a disintegrin and metalloproteinase doma	mAb & diag & s.m.	plasma membrane
	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	mAb & diag	plasma membrane
	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	diag	secreted
	451668	Z43948	Hs.326444	cartilage acidic protein 1	mAb & diag	plasma membrane
75	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	diag	
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	diag	extracellular
	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR	mAb	plasma membrane
	452838	U65011	Hs.30743	preferentially expressed antigen in mela	CTL	nuclear
	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso	CTL & s.m.	nuclear
80	457489	AI693815	Hs.127179	cryptic gene	diag	secreted

TABLE 14B

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey CAT Number Accession



5	414883	15024_1	AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245 AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW960304 T29812 AA476873 BE297387 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 AI719075 BE270172 BE269819 AA889955 AI204630 W25243 AI935150 AA872039 W72395 T99630 AI422691 H98460 N31428 BE255916 H03265 AI857576 AA776920 AA910644 AA459522 AA293140 AW514667 R75953 AW662396 AA662522 AI865147 AI423153 AW262230 AA584410 AA583187 AW024595 AW069734 AI828996 AA282997 AA876046 AW613002 AA527373 AW972459 AI831360 AA621337 AA100926 AA772418 AA594628 AI033892 W95096 AI034317 AA398727 AI085031 N95210 AI459432 AI041437 AA932124 AA627684 AA935829 AI004827 AI423513 AI094597 H42079 R54703 AI630359 AA617681 AA978045 AA643280 W44561 AI991988 AI537692 AI090262 AA740817 AI312104 AI911822 AA416871 AI185409 AA129784 AA701623 AI075239 AI139549 AA633648 AI339996 AI336880 AA399239 AI078708 AI085351 AI362835 AI346618 AI146955 AI989380 AI348243 N92892 AA765850 AI494230 AI278887 AA962596 AI492600 W80435 AA001979 R97424 AI129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785 AI494211 AW059601 AW886710 R92790 N59755 AI361128 AW589407 H47725 H97534 H48076 H48450 T99631 AW300758 H03431 R76789 AA954344 H77576 R96823 AI457100 N92845 N49682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923 H03266 BE261919 AA769633 AA480310 AA507454 AA910586 AI203723 AW104725 W25611 W25071 T88980 H03513 T77589 R99156 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283672 1545037583327_1AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532 AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AI161014 AA099554 R69067
---	--------	---------	--

20	TABLE 14C		
25	Pkey:	Unique number corresponding to an Eos probeset	
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.	
	Strand:	Indicates DNA strand from which exons were predicted.	
	Nt_position:	Indicates nucleotide positions of predicted exons.	
	Pkey	Ref	Strand Nt_position
30	402075	8117407	Plus 121907-122035,122804-122921,124019-124161,124455-124610,125672-126076



TABLE 15A: Information for all sequences in Table 16

Table 15A shows the Seq ID No, Pkey, ExAccn, UnigenelD, and Unigene Title for all of the sequences in Table 16.

Table 15B show the accession numbers for those Pkey's lacking UnigenelD's for table 15A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 15C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 15A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Seq ID No: Sequence ID number  
 Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Unigene Title: Unigene gene title

Seq ID No:	Pkey	ExAccn	UnigenelD	Unigene Title
Seq ID No: 1 & 2	410407	X66839	Hs.63287	carbonic anhydrase IX
Seq ID No: 3 & 4	412719	AW016610	Hs.816	ESTs
Seq ID No: 5 & 6	417034	NM_006183	Hs.80962	neurotensin
Seq ID No: 7 & 8	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
Seq ID No: 9 & 10	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 11 & 12	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 13 & 14	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 15 & 16	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
Seq ID No: 17 & 18	439285	AL133916		hypothetical protein FLJ20093
Seq ID No: 19 & 20	413753	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalinin
Seq ID No: 21 & 22	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog
Seq ID No: 23 & 24	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
Seq ID No: 25 & 26	412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkines
Seq ID No: 27 & 28	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
Seq ID No: 29 & 30	452838	U65011	Hs.30743	preferentially expressed antigen in mela
Seq ID No: 31 & 32	418663	AK001100	Hs.41690	desmocollin 3
Seq ID No: 33 & 34	418663	AK001100	Hs.41690	desmocollin 3
Seq ID No: 35 & 36	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibito
Seq ID No: 37 & 38	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
Seq ID No: 39 & 40	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
Seq ID No: 41 & 42	431846	BE019924	Hs.271580	uroplakin 1B
Seq ID No: 43 & 44	418830	BE513731	Hs.88959	hypothetical protein MGC4816
Seq ID No: 45 & 46	424098	AF077374	Hs.139322	small proline-rich protein 3
Seq ID No: 47 & 48	443648	AI085377	Hs.143610	ESTs
Seq ID No: 49	311034	BE567130	Hs.311389	ESTs, Highly similar to NKGD_HUMAN NKG2-
Seq ID No: 50 & 51	408522	AI541214	Hs.46320	Small proline-rich protein SPRK [human,
Seq ID No: 52 & 53	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
Seq ID No: 54 & 55	435505	AF200492	Hs.211238	interleukin-1 homolog 1
Seq ID No: 56 & 57	417366	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)
Seq ID No: 58 & 59	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
Seq ID No: 60 & 61	441020	W79283	Hs.35962	ESTs
Seq ID No: 62 & 63	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys
Seq ID No: 64 & 65	429538	BE182592	Hs.11261	small proline-rich protein 2A
Seq ID No: 66 & 67	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte
Seq ID No: 68 & 69	444371	BE540274	Hs.239	forkhead box M1
Seq ID No: 70 & 71	444371	BE540274	Hs.239	forkhead box M1
Seq ID No: 72 & 73	444371	BE540274	Hs.239	forkhead box M1
Seq ID No: 74 & 75	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
Seq ID No: 76 & 77	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
Seq ID No: 78 & 79	429259	AA420450	Hs.292911	Plakophilin
Seq ID No: 80 & 81	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu
Seq ID No: 82 & 83	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
Seq ID No: 84 & 85	423662	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro
Seq ID No: 86 & 87	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
Seq ID No: 88 & 89	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
Seq ID No: 90 & 91	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor
Seq ID No: 92 & 93	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr
Seq ID No: 94 & 95	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated
Seq ID No: 96 & 97	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re
Seq ID No: 98 & 99	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic
Seq ID No: 100 & 101	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o
Seq ID No: 102 & 103	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
Seq ID No: 104 & 105	409103	AF251237	Hs.112208	XAGE-1 protein
Seq ID No: 106 & 107	417542	J04129	Hs.82269	progesterone-associated endometrial prote
Seq ID No: 108 & 109	428471	X57348	Hs.184510	stratifin
Seq ID No: 110 & 111	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member
Seq ID No: 112 & 113	414761	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2
Seq ID No: 114 & 115	418203	X54942	Hs.83758	CDC28 protein kinase 2
Seq ID No: 116	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m
Seq ID No: 117 & 118	437016	AU076916	Hs.5398	guanine monophosphate synthetase
Seq ID No: 119 & 120	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
Seq ID No: 121 & 122	446989	AK001898	Hs.16740	hypothetical protein FLJ11036
Seq ID No: 123 & 124	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein
Seq ID No: 125 & 126	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B



5	Seq ID No: 127 & 128	414430	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1
	Seq ID No: 129 & 130	418462	BE001596	Hs.85266	integrin, beta 4
	Seq ID No: 131 & 132	100668	L05424	Hs.169610	CD44 antigen (homing function and Indian
	Seq ID No: 133 & 134	458933	AI638429	Hs.24763	RAN binding protein 1
	Seq ID No: 135 & 136	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
10	Seq ID No: 137 & 138	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 139 & 140	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 141 & 142	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 143 & 144	446269	AW263155	Hs.14559	hypothetical protein FLJ10540
	Seq ID No: 145 & 146	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur
15	Seq ID No: 147 & 148	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con
	Seq ID No: 149 & 150	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma
	Seq ID No: 151 & 152	439606	W79123	Hs.58561	G protein-coupled receptor 87
	Seq ID No: 153 & 154	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 155 & 156	453884	AA355925	Hs.36232	KIAA0186 gene product
20	Seq ID No: 157 & 158	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 159 & 160	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 161 & 162	404877			NM_005365:Homo sapiens melanoma antigen,
	Seq ID No: 163 & 164	413129	AF292100	Hs.104613	RP42 homolog
	Seq ID No: 165 & 166	413281	AA861271	Hs.222024	transcription factor BMAL2
25	Seq ID No: 167 & 168	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 169 & 170	416819	U77735	Hs.80205	pim-2 oncogene
	Seq ID No: 171 & 172	451320	AW118072		diacylglycerol kinase, zeta (104kD)
	Seq ID No: 173 & 174	418543	NM_005329	Hs.85962	hyaluronan synthase 3
	Seq ID No: 175 & 176	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member
30	Seq ID No: 177 & 178	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)
	Seq ID No: 179 & 180	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 181 & 182	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 183 & 184	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 185 & 186	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
35	Seq ID No: 187 & 188	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 189 & 190	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 191 & 192	448993	AI471630	Hs.8127	KIAA0144 gene product
	Seq ID No: 193 & 194	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 195 & 196	430393	BE185030	Hs.241305	estrogen-responsive B box protein
40	Seq ID No: 197 & 198	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol
	Seq ID No: 199 & 200	420462	AF050147	Hs.97932	chondromodulin I precursor
	Seq ID No: 201 & 202	102963	X02404	Hs.274534	calcitonin-related polypeptide, beta
	Seq ID No: 203 & 204	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid
	Seq ID No: 205 & 206	101175	U82671	Hs.36980	melanoma antigen, family A, 2
45	Seq ID No: 207 & 208	429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like
	Seq ID No: 209 & 210	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 211 & 212	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 213 & 214	131927	AJ003112	Hs.34780	doublecortin; lissencephaly, X-linked (d
	Seq ID No: 215 & 216	428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT
50	Seq ID No: 217 & 218	427335	AA448542	Hs.251677	G antigen 7B
	Seq ID No: 219 & 220	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini
	Seq ID No: 221 & 222	114346	AL137256	Hs.130489	ATPase, aminophospholipid transporter-II
	Seq ID No: 223 & 224	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9
	Seq ID No: 225 & 226	404440			NM_021048:Homo sapiens melanoma antigen,
55	Seq ID No: 227 & 228	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibito
	Seq ID No: 229 & 230	103312	Y12642	Hs.3185	lysosomal
	Seq ID No: 231 & 232	320843	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr
	Seq ID No: 233	429065	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT
	Seq ID No: 234 & 235	446102	AW168067	Hs.317694	ESTs
60	Seq ID No: 236 & 237	330495	U47924	Hs.71642	guanine nucleotide binding protein (G pr
	Seq ID No: 238	413573	AI733859	Hs.149089	ESTs
	Seq ID No: 239 & 240	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
	Seq ID No: 241 & 242	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
	Seq ID No: 243 & 244	332180	AF134160	Hs.7327	claudin 1
65	Seq ID No: 245	437915	AI637993	Hs.202312	Homo sapiens clone N11 NTera2D1 teratoca
	Seq ID No: 246 & 247	441553	AA281219	Hs.121296	ESTs
	Seq ID No: 248 & 249	331692	AI683487	Hs.152213	wingless-type MMTV integration site fami
	Seq ID No: 250 & 251	429413	NM_014058	Hs.201877	DESC1 protein
	Seq ID No: 252 & 253	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis
70	Seq ID No: 254 & 255	448357	N20169	Hs.108923	RAB38, member RAS oncogene family
	Seq ID No: 256 & 257	446292	AF081497	Hs.279682	Rh type C glycoprotein
	Seq ID No: 258 & 259	416209	AA236776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h
	Seq ID No: 260 & 261	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1
	Seq ID No: 262 & 263	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito
75	Seq ID No: 264 & 265	439223	AW238299	Hs.250618	UL16 binding protein 2
	Seq ID No: 266 & 267	429228	AI553633	Hs.326447	ESTs
	Seq ID No: 268 & 269	409757	NM_001898	Hs.123114	cystatin SN
	Seq ID No: 270 & 271	411089	AA456454	Hs.214291	cell division cycle 2-like 1 (PITSLRE pr
	Seq ID No: 272 & 273	436511	AA721252	Hs.291502	ESTs
80	Seq ID No: 274 & 275	428969	AF120274	Hs.194689	artemin
	Seq ID No: 276 & 277	428969	AF120274	Hs.194689	artemin
	Seq ID No: 278 & 279	428969	AF120274	Hs.194689	artemin
	Seq ID No: 280 & 281	428969	AF120274	Hs.194689	artemin
	Seq ID No: 282	407137	T97307		gb:ye53h05.s1 Soares fetal liver spleen
85	Seq ID No: 283 & 284	412723	AA648459	Hs.335951	hypothetical protein AF301222
	Seq ID No: 285 & 286	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 287 & 288	405770			NM_002362:Homo sapiens melanoma antigen,
	Seq ID No: 289 & 290	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13
	Seq ID No: 291 & 292	414774	X02419	Hs.77274	plasminogen activator, urokinase



5	Seq ID No: 293 & 294	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub
	Seq ID No: 295 & 296	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 297 & 298	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 299 & 300	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 301 & 302	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
10	Seq ID No: 303 & 304	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 305 & 306	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso
	Seq ID No: 307 & 308	403478			NM_022342:Homo sapiens kinesin protein 9
	Seq ID No: 309	441525	AW241867	Hs.127728	ESTs
	Seq ID No: 310 & 311	434105	AW952124	Hs.13094	presenilins associated rhomboid-like pro
15	Seq ID No: 312 & 313	428810	AF068236	Hs.193788	nitric oxide synthase 2A (inducible, hep
	Seq ID No: 314 & 315	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B
	Seq ID No: 316 & 317	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f
	Seq ID No: 318 & 319	409228	R16811	Hs.22010	ESTs, Weakly similar to 2109260A B cell
	Seq ID No: 320 & 321	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg
20	Seq ID No: 322 & 323	413582	AW295647	Hs.71331	hypothetical protein MGC5350
	Seq ID No: 324 & 325	438403	AA806607	Hs.292206	ESTs
	Seq ID No: 326 & 327	403329			unnamed protein product [Homo sapiens]
	Seq ID No: 328 & 329	409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.
	Seq ID No: 330 & 331	119073	BE245360	Hs.279477	v-ets erythroblastosis virus E26 oncogen
25	Seq ID No: 332 & 333	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	Seq ID No: 334 & 335	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	Seq ID No: 336 & 337	101345	NM_005795	Hs.152175	calcitonin receptor-like
	Seq ID No: 338 & 339	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	Seq ID No: 340 & 341	102012	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
30	Seq ID No: 342 & 343	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	Seq ID No: 344 & 345	134299	AW580939	Hs.97199	complement component C1q receptor
	Seq ID No: 346 & 347	412719	AW016610	Hs.816	ESTs
	Seq ID No: 348 & 349	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
	Seq ID No: 350 & 351	128924	BE279383	Hs.26557	plakophilin 3
35	Seq ID No: 352 & 353	100486	T19006	Hs.10842	RAN, member RAS oncogene family
	Seq ID No: 354 & 355	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 356 & 357	409459	D86407	Hs.54481	low density lipoprotein receptor-related
	Seq ID No: 358 & 359	330493	M27826		endogenous retroviral protease
	Seq ID No: 360 & 361	417866	AW067903	Hs.82772	collagen, type XI, alpha 1
40	Seq ID No: 362 & 363	418113	AI272141	Hs.83484	SRY (sex determining region Y)-box 4
	Seq ID No: 364 & 365	437016	AU076916	Hs.5398	guanine monophosphate synthetase
	Seq ID No: 366 & 367	429612	AF062649	Hs.252587	pituitary tumor-transforming 1
	Seq ID No: 368 & 369	440704	M69241	Hs.162	insulin-like growth factor binding prote
	Seq ID No: 370 & 371	431221	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye
45	Seq ID No: 372 & 373	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 374 & 375	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 376 & 377	132354	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)
	Seq ID No: 378 & 379	424441	X14850	Hs.147097	H2A histone family, member X
	Seq ID No: 380 & 381	103768	AF086009	Hs.296398	gb:Homo sapiens full length insert cDNA
50	Seq ID No: 382 & 383	417512	X76534	Hs.82226	glycoprotein (transmembrane) nmb
	Seq ID No: 384 & 385	425266	J00077	Hs.155421	alpha-fetoprotein
	Seq ID No: 386 & 387	424503	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,
	Seq ID No: 388 & 389	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	Seq ID No: 390 & 391	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
55	Seq ID No: 392 & 393	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	Seq ID No: 394 & 395	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	Seq ID No: 396 & 397	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2
	Seq ID No: 398 & 399	418506	AA084248	Hs.85339	G protein-coupled receptor 39
	Seq ID No: 400 & 401	423961	D13666	Hs.136348	periostin (OSF-2os)
60	Seq ID No: 402 & 403	414812	X72755	Hs.77367	monokine induced by gamma interferon
	Seq ID No: 404 & 405	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
	Seq ID No: 406 & 407	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
	Seq ID No: 408 & 409	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse
	Seq ID No: 410 & 411	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy
65	Seq ID No: 412 & 413	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti
	Seq ID No: 414 & 415	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated
	Seq ID No: 416 & 417	411789	AF245505	Hs.72157	Adlican
	Seq ID No: 418 & 419	428698	AA852773	Hs.334838	KIAA1866 protein
	Seq ID No: 420 & 421	450098	W27249	Hs.8109	hypothetical protein FLJ21080
70	Seq ID No: 422 & 423	421552	AF026692	Hs.105700	secreted frizzled-related protein 4
	Seq ID No: 424 & 425	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR
	Seq ID No: 426 & 427	450375	AA009647		a disintegrin and metalloproteinase doma
	Seq ID No: 428 & 429	426215	AW963419	Hs.155223	stanniocalcin 2
	Seq ID No: 430 & 431	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin
75	Seq ID No: 432 & 433	432201	AI538613	Hs.298241	Transmembrane protease, serine 3
	Seq ID No: 434 & 435	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph
	Seq ID No: 436 & 437	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447
	Seq ID No: 438 & 439	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn
	Seq ID No: 440 & 441	447033	AI357412	Hs.157601	ESTs
80	Seq ID No: 442 & 443	447033	AI357412	Hs.157601	ESTs
	Seq ID No: 444 & 445	447033	AI357412	Hs.157601	ESTs
	Seq ID No: 446 & 447	115522	BE614387	Hs.333893	c-Myc target JPO1
	Seq ID No: 448 & 449	410418	D31382	Hs.63325	transmembrane protease, serine 4
	Seq ID No: 450 & 451	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
85	Seq ID No: 452 & 453	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
	Seq ID No: 454 & 455	452461	N78223	Hs.108106	transcription factor
	Seq ID No: 456 & 457	412420	AL035668	Hs.73853	bone morphogenetic protein 2
	Seq ID No: 458 & 459	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara
	Seq ID No: 460 & 461	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagone



	Seq ID No: 462 & 463	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 [H.s.a
	Seq ID No: 464 & 465	402075			ENSP00000251056*:Plasma membrane calcium
	Seq ID No: 466 & 467	421110	AJ250717	Hs.1355	cathepsin E
5	Seq ID No: 468 & 469	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 470 & 471	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 472 & 473	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 474 & 475	422282	AF019225	Hs.114309	apolipoprotein L
	Seq ID No: 476 & 477	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member
10	Seq ID No: 478 & 479	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),
	Seq ID No: 480 & 481	427747	AW411425	Hs.180655	serine/threonine kinase 12
	Seq ID No: 482 & 483	420281	AI623693	Hs.323494	Predicted cation efflux pump
	Seq ID No: 484 & 485	405932			C15000305:gil3806122[gb]AAC69198.1) (AF0
	Seq ID No: 486 & 487	405932			C15000305:gil3806122[gb]AAC69198.1) (AF0
15	Seq ID No: 488 & 489	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane
	Seq ID No: 490 & 491	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy
	Seq ID No: 492 & 493	417079	U65590	Hs.81134	interleukin 1 receptor antagonist
	Seq ID No: 494 & 495	430890	X54232	Hs.2699	glypican 1
	Seq ID No: 496 & 497	419721	NM_001650	Hs.288650	aquaporin 4
20	Seq ID No: 498 & 499	444471	AB020684	Hs.11217	KIAA0877 protein
	Seq ID No: 500 & 501	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote
	Seq ID No: 502 & 503	433800	AI034361	Hs.135150	lung type-I cell membrane-associated gly
	Seq ID No: 504 & 505	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
	Seq ID No: 506 & 507	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
25	Seq ID No: 508 & 509	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte
	Seq ID No: 510 & 511	410407	X66839	Hs.63287	carbonic anhydrase IX
	Seq ID No: 512 & 513	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s
	Seq ID No: 514 & 515	412719	AW016610	Hs.816	ESTs
	Seq ID No: 516 & 517	417034	NM_006183	Hs.80962	neurotensin
30	Seq ID No: 518 & 519	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
	Seq ID No: 520 & 521	413753	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalinin
	Seq ID No: 522 & 523	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
	Seq ID No: 524 & 525	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
	Seq ID No: 526 & 527	418663	AK001100	Hs.41690	desmocollin 3
35	Seq ID No: 528 & 529	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 530 & 531	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
	Seq ID No: 532 & 533	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
	Seq ID No: 534 & 535	431846	BE019924	Hs.271580	uroplakin 1B
	Seq ID No: 536 & 537	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
40	Seq ID No: 538 & 539	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
	Seq ID No: 540 & 541	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
	Seq ID No: 542 & 543	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
	Seq ID No: 544 & 545	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
	Seq ID No: 546 & 547	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor
45	Seq ID No: 548 & 549	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
	Seq ID No: 550 & 551	417542	J04129	Hs.82269	progesterone-associated endometrial prote
	Seq ID No: 552 & 553	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
	Seq ID No: 554 & 555	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
	Seq ID No: 556 & 557	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
50	Seq ID No: 558 & 559	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B
	Seq ID No: 560 & 561	418462	BE001596	Hs.85266	integrin, beta 4
	Seq ID No: 562 & 563	410274	AA381807	Hs.61762	hypoxia-inducible protein 2
	Seq ID No: 564 & 565	439606	W79123	Hs.58561	G protein-coupled receptor 87
	Seq ID No: 566 & 567	404877			NM_005365:Homo sapiens melanoma antigen,
55	Seq ID No: 568 & 569	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 570 & 571	418543	NM_005329	Hs.85962	hyaluronan synthase 3
	Seq ID No: 572 & 573	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 574 & 575	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 576 & 577	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
60	Seq ID No: 578 & 579	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 580 & 581	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 582 & 583	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 584 & 585	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 586 & 587	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
65	Seq ID No: 588 & 589	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 590 & 591	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini
	Seq ID No: 592 & 593	332180	AF134160	Hs.7327	claudin 1
	Seq ID No: 594 & 595	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
	Seq ID No: 596 & 597	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
70	Seq ID No: 598 & 599	439223	AW238299	Hs.250618	UL16 binding protein 2
	Seq ID No: 600 & 601	409757	NM_001898	Hs.123114	cystatin SN
	Seq ID No: 602 & 603	428969	AF120274	Hs.194689	artemin
	Seq ID No: 604 & 605	428969	AF120274	Hs.194689	artemin
	Seq ID No: 606 & 607	428969	AF120274	Hs.194689	artemin
75	Seq ID No: 608 & 609	428969	AF120274	Hs.194689	artemin
	Seq ID No: 610 & 611	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 612 & 613	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 614 & 615	414774	X02419	Hs.77274	plasminogen activator, urokinase
	Seq ID No: 616 & 617	407944	R34008	Hs.239727	desmocollin 2
80	Seq ID No: 618 & 619	407944	R34008	Hs.239727	desmocollin 2
	Seq ID No: 620 & 621	457489	AI693815	Hs.127179	cryptic gene
	Seq ID No: 622 & 623	429547	AW009166	Hs.99376	ESTs
	Seq ID No: 624 & 625	407242	M18728		gb:Human nonspecific crossreacting antig
	Seq ID No: 626 & 627	407242	M18728		gb:Human nonspecific crossreacting antig
85	Seq ID No: 628 & 629	407242	M18728		gb:Human nonspecific crossreacting antig
	Seq ID No: 630 & 631	444006	BE395085	Hs.10086	type I transmembrane protein Fn14



5	Seq ID No: 632 & 633	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma
	Seq ID No: 634 & 635	422109	S73265	Hs.1473	gastrin-releasing peptide
	Seq ID No: 636 & 637	419235	AW470411	Hs.288433	neurotrimin
	Seq ID No: 638 & 639	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc
	Seq ID No: 640 & 641	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy
10	Seq ID No: 642 & 643	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu
	Seq ID No: 644 & 645	448243	AW369771	Hs.52620	integrin, beta 8
	Seq ID No: 646 & 647	426427	M86699	Hs.169840	TTK protein kinase
	Seq ID No: 648 & 649	445537	AJ245671	Hs.12844	EGF-like-domain, multiple 6
	Seq ID No: 650 & 651	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6
15	Seq ID No: 652 & 653	428450	NM_014791	Hs.184339	KIAA0175 gene product
	Seq ID No: 654 & 655	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,
	Seq ID No: 656 & 657	453392	U23752	Hs.32964	SRV (sex determining region Y)-box 11
	Seq ID No: 658 & 659	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic
	Seq ID No: 660 & 661	425776	U25128	Hs.159499	parathyroid hormone receptor 2
20	Seq ID No: 662 & 663	425776	U25128	Hs.159499	parathyroid hormone receptor 2
	Seq ID No: 664 & 665	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha
	Seq ID No: 666 & 667	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7
	Seq ID No: 668 & 669	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
	Seq ID No: 670 & 671	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
25	Seq ID No: 672 & 673	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
	Seq ID No: 674 & 675	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
	Seq ID No: 676 & 677	410001	AB041036	Hs.57771	kallikrein 11
	Seq ID No: 678 & 679	426501	AW043782	Hs.293616	ESTs
	Seq ID No: 680 & 681	408369	R38438	Hs.182575	solute carrier family 15 (H??? transport
30	Seq ID No: 682 & 683	445413	AA151342	Hs.12677	CGI-147 protein
	Seq ID No: 684 & 685	422424	AI186431	Hs.296638	prostate differentiation factor
	Seq ID No: 686 & 687	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,
	Seq ID No: 688 & 689	420610	AI683183	Hs.99348	distal-less homeo box 5

TABLE 15B

35  
 Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

40	Pkey	CAT Number	Accession
	309931	AW341683	
	330493	33264_5	M27826 R78416 AA307645 AW957879 AW957800 AA633529 H03662
45	439285	47065_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882
			AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077
	450375	83327_1	AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532
50			AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AI161014 AA099554 R69067
	451320	86576_1	AW118072 AI631982 T15734 AA224195 AI701458 W20198 F26326 AA890570 N90552 AW071907 AI671352 AI375892 T03517 R88265
			AI124088 AA224388 AI084316 AI354686 T33652 AI140719 AI720211 T03490 AI372637 T15415 AW205836 AA630384 T03515 T33230
			AA017131 AA443303 T33623 AI222556 T33511 T33785 AI419606 D55612

TABLE 15C

55  
 Pkey: Unique number corresponding to an Eos probeset  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 NL\_position: Indicates nucleotide positions of predicted exons.

60	Pkey	Ref	Strand	NL_position
	402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
	403329	8516120	Plus	96450-96598
65	403478	9958258	Plus	116458-116564
	404440	7528051	Plus	80430-81581
	404877	1519284	Plus	1095-2107
	405770	2735037	Plus	61057-62075
	405932	7767812	Minus	123525-123713



Table 16

Seq ID NO: 1 DNA sequence  
Nucleic Acid Accession #: NM\_001216  
Coding sequence: 43..1422

```

1      11      21      31      41      51
|      |      |      |      |      |
10  G C C C G T A C A C   A C C G T G T G C T   G G G A C A C C C C   A C A G T C A G C C   G C A T G G C T C C   C C T G T G C C C C   60
    A G C C C C T G G C   T C C C T C T G T T   G A T C C C G G C C   C C T G C T C C A G   G C C T C A C T G T   G C A A C T G C T G   120
    C T G T C A C T G C   T G C T T C T G A T   G C C T G T C C A T   C C C C A G A G G T   T G C C C C G G A T   G C A G G A G G A T   180
    T C C C C C T T G G   G A G G A G G C T C   T T C T G G G G A A   G A T G A C C C A C   T G G G C G A G G A   G G A T C T G C C C   240
    A G T G A A G A G G   A T T C A C C C A G   A G A G G A G G A T   C C A C C C G G A G   A G G A G G A T C T   A C C T G G A G A G   300
    G A G G A T C T A C   C T G G A G A G G A   G G A T C T A C C T   G A A G T T A A G C   C T A A A T C A G A   A G A A G A G G G C   360
15  T C C C T G A A G T   T A G A G G A T C T   A C C T A C T G T T   G A G G C T C C T G   G A G A T C C T C A   A G A A C C C C A G   420
    A A T A A T G C C C   A C A G G G A C A A   A G A A G G G G A T   G A C C A G A G T C   A T T G G C G C T A   T G G A G G C G A C   480
    C C G C C C T G G C   C C C G G G T G T C   C C C A G C C T G C   G C G G G C C G C T   T C C A G T C C C C   G G T G G A T A T C   540
    C G C C C C A G C   T C G C C G C C T T   C T G C C C G G C C   C T G C G C C C C C   T G G A A C T C C T   G G G C T T C C A G   600
    C T C C C G C C G C   T C C C A G A A C T   G C G C C T G C G C   A A C A A T G G C C   A C A G T G T G C A   A C T G A C C C T G   660
20  C C T C C T G G G C   T A G A G A T G G C   T C T G G G T C C C   G G G C G G G A G T   A C C G G G C T C T   G C A G C T G C A T   720
    C T G C A C T G G G   G G G C T G C A G G   T C G T C C G G G C   T C G T G G A A G G   C C A C C G T T T C   C A C C C G T T T C   780
    C C T G C C G A G A   T C C A C G T G G T   T C A C C T C A G C   A C C G C C T T T G   C C A G A G T T G A   C G A G G C C T T G   840
    G G G C G C C C G G   G A G G C C T G G C   C G T G T T G G C C   G C C T T T C T G G   A G G A G G G C C C   G G A A G A A A A C   900
25  A G T G C C T A T G   A G C A G T T G C T   G T C T C G C T T G   G A A G A A A T C G   C T G A G G A A G G   C T C A G A G A C T   960
    C A G G T C C C A G   G A C T G G A C A T   A T C T G C A C T C   C T G C C C T C T G   A C T T C A G C C G   C T A C T T C C A A   1020
    T A T G A G G G G T   C T C T G A C T A C   A C C G C C C T G T   G C C C A G G G T G   T C A T C T G G A C   T G T G T T T A A C   1080
    C A G A C A G T G A   T G C T G A G T G C   T A A G C A G C T C   C A C A C C C T C T   C T G A C A C C C T   G T G G G G A C C T   1140
    G G T G A C T C T C   G G C T A C A G C T   G A A C T T C C G A   G C G A C G C A G C   C T T T G A A T G G   G C G A G T G A T T   1200
30  G A G G C C T C C T   T C C C T G C T G G   A G T G G A C A G C   A G T C C T C G G G   C T G C T G A G C C   A G T C C A G C T G   1260
    A A T T C C T G C C   T G G C T G C T G G   T G A C A T C C T A   G C C C T G G T T T   T T G G C C T C C T   T T T T G C T G T C   1320
    A C C A G C G T C G   C G T T C C T T G T   G C A G A T G A G A   A G G C A G C A C A   G A A G G G G A A C   C A A A G G G G G T   1380
    G T G A G C T A C C   G C C C A G C A G A   G G T A G C C G A G   A C T G G A G C C T   A G A G G C T G G A   T C T T G G A G A A   1440
    T G T G A G A A G C   C A G C C A G A G G   C A T C T G A G G G   G G A G C C G G T A   A C T G T C C T G T   C C T G C T C A T T   1500
35  A T G C C A C T T C   C T T T T A A C T G   C C A A G A A A T T   T T T T A A A A T A   A A T A T T T A T A   A T

```

Seq ID NO: 2 Protein sequence:  
Protein Accession #: NP\_001207

```

1      11      21      31      41      51
|      |      |      |      |      |
40  M A P L C P S P W L   P L L I P A P A P G   L T V Q L L L S L L   L L M P V H P Q R L   P R M Q E D S P L G   G G S S G E D D P L   60
    G E E D L P S E E D   S P R E E D P P G E   E D L P G E E D L P   G E E D L P E V K P   K S E E G S L K L   E D L P T V E A P G   120
    D P Q E P Q N N A H   R D K E G D D Q S H   W R Y G G D P P W P   R V S P A C A G R F   Q S P V D I R P Q L   A A F C P A L R P L   180
45  E L L G F Q L P P L   P E L R L R N N G H   S V Q L T L P P G L   E M A L G P G R E Y   R A L Q L H L H W G   A A G R P G S E H T   240
    V E G H R F P A E I   H V V H L S T A F A   R V D E A L G R P G   G L A V L A A F L E   E G P E E N S A Y E   Q L L S R L E E I A   300
    E E G S E T Q V P G   L D I S A L L P S D   F S R Y F Q Y E G S   L T T P P C A Q G V   I W T V F N Q T V M   L S A K Q L H T L S   360
    D T L W G P G D S R   L Q L N F R A T Q P   L N G R V I E A S F   P A G V D S S P R A   A E P V Q L N S C L   A A G D I L A L V F   420
    G L L F A V T S V A   F L V Q M R R Q H R   R G T K G G V S Y R   P A E V A E T G A

```

Seq ID NO: 3 DNA sequence  
Nucleic Acid Accession #: BC013923  
Coding sequence: 438-1391

```

1      11      21      31      41      51
|      |      |      |      |      |
55  A G C G G G G T T G   T C T A T T A A C T   T G T T C A A A A A   G T A T C A G G A G   T T G T C A A G G C   A G A G A A G A G A   60
    G T G T T T G C A A   A A G G G G G A A A   G T A G T T T G C T   G C C T C T T T A A   G A C T A G G A C T   G A G A G A A A G A   120
    A G A G G A G A G A   G A A A G A A A G G   G A G A G A A G T T   T G A G C C C C A G   G C T T A A G C C T   T T C C A A A A A A   180
60  T A A T A A T A A C   A A T C A T C G G C   G G C G G C A G G A   T C G G C C A G A G   G A G G A G G G A A   G C G C T T T T T T   240
    T G A T C C T G A T   T C C A G T T T G C   C T C T C T C T T T   T T T T C C C C C A   A A T T A T T C T T   C G C C T G A T T T   300
    T C C T C G C G G A   G C C C T G C G C T   C C C G A C A C C C   C C G C C C G C C T   C C C C T C C T C C   T C T C C C C C C G   360
    C C C G C G G G C C   C C C C A A A G T C   C C G G C C G G G C   C G A G G G T C G G   C G G C C G C C G G   C G G G C C G G G C   420
    C C G C G C A C A G   C G C C C G C A T G   T A C A A C A T G A   T G G A G A C G G A   G C T G A A G C C G   C C G G G C C C G C   480
65  A G C A A A C T T C   G G G G G G C G G C   G G C G G C A A C T   C C A C C G C G G C   G G C G G C C G G C   G G C A A C C A G A   540
    A A A A C A G C C C   G G A C C G C G T C   A A G C G G C C C A   T G A A T G C C T T   C A T G G T G T G G   T C C C G C G G G C   600
    A G C G G C G C A A   G A T G G C C C A G   A G A A C C C C A A   A G A T G C A C A A   C T C G G A G A T C   A G C A A G C G C C   660
    T G G G C G C C G A   G T G G A A A C T T   T T G T C G G A G A   C G G A G A A G C G   G C C G T T C A T C   G A C G A G G C T A   720
    A G C G G C T G C G   A G C G C T G C A C   A T G A A G G A G C   A C C C G G A T T A   T A A A T A C C G G   C C C C G G C G G A   780
70  A A A C C A A G A C   G C T C A T G A A G   A A G G A T A A G T   A C A C G T G C C   C G G C G G G C T G   C T G G C C C C C G   840
    G C G G C A A T A G   C A T G G C G A G C   G G G G T C G G G G   T G G G C G C C G G   C C T G G G C G C G   G G C G T G A A C C   900
    A G C G C A T G G A   C A G T T A C G C G   C A C A T G A A C G   G C T G G A G C A A   C G G C A G C T A C   A G C A T G A T G C   960
    A G G A C C A G C T   G G G C T A C C C G   C A G C A C C C G G   G C C T C A A T G C   G C A C G G C G C A   G C G C A G A T G C   1020
    A G C C C A T G C A   C C G C T A C G A C   G T G A G C G C C C   T G C A G T A C A A   C T C C A T G A C C   A G C T C G C A G A   1080
75  C C T A C A T G A A   C G G C T C G C C C   A C C T A C A G C A   T G T C C T A C T C   G C A G C A G G G C   A C C C C T G G C A   1140
    T G G C T C T T G G   C T C C A T G G G T   T C G G T G G T C A   A G T C C G A G G C   C A G C T C C A G C   C C C C C T G T G G   1200
    T T A C C T C T T C   C T C C C A C T C C   A G G G C G C C C T   G C C A G G C C G G   G G A C C T C C G G   G A C A T G A T C A   1260
    G C A T G T A T C T   C C C C G G C G C C   G A G G T G C C G G   A A C C C G C C G C   C C C C A G C A G A   C T T C A C A T G T   1320
    C C C A G C A C T A   C C A G A G C G G C   C C G G T G C C C G   G C A C G G C C A T   T A A C G G C A C A   C T G C C C C T C T   1380
    C A C A C A T G T G   A G G C C C G G A C   A G C G A A C T G G   A G G G G G G A G A   A A T T T T C A A A   G A A A A C G A G   1440
80  G G A A A T G G G A   G G G G T G C A A A   A G A G G A G A G T   A A G A A A C A G C   A T G G A G A A A A   C C C G G T A C G C   1500
    T C A A A A A A A A   A A A A A A A A A A   A A A A T C C C A T   C A C C C A C A G C   A A A T G A C A G C   T G C A A A A G A G   1560
    A A C A C C A A T C   C C A T C C A C A C   T C A C G C A A A A   A C C G C G A T G C   C G A C A A G A A A   A C T T T T A T G A   1620
    G A G A G A T C C T   G G A C T T C T T T   T K G G G G A C T   A T T T T T G T A C   A G A G A A A A C C   T G G G G A G G G T   1680
85  G G G G A G G G G C   G G G G A A T G G A   C C T T G T A T A G   A C T C T G G A G G A   A A G A A A G C T A   C G A A A A A C T T   1740
    T T T A A A A G T T   C T A G T G G T A C   G G T A G G A G C T   T T G C A G G A A G   T T T G C A A A A G   T C T T T A C C A A   1800
    T A A T A T T T A G   A G C T A G T C T C   C A A G C G A C G A   A A A A A A T G T T   T T A A T A T T T G   C A A G C A A C T T   1860
    T T G T A C A G T A   T T T A T C G A G A   T A A A C A T G G C   A A T C A A A A T G   T C C A T T G T T T   A T A A G C T G A G   1920

```



AATTTGCCAA TATTTTTCAT GGAGAGGCTT CTTGCTGAAT TTTGATTCTG CAGCTGAAAT 1980  
 TTAGGACAGT TGCAAACGTG AAAAGAAGAA AATTATTCAA ATTTGGACAT TTTAATTGTT 2040  
 TAAAAATTGT ACAAAGGAA AAAATTAGAA TAAGTACTGG CGAACCATCT CTGTGGTCTT 2100  
 5 GTTTAAAAAG GGCAAAAGTT TTAGACTGTA CTAAATTTTA TAACTTACTG TTTAAAGCAA 2160  
 AAATGGCCAT GCAGGTTGAC ACCGTTGGTA ATTTATAATA GCTTTTGTTC GATCCCAACT 2220  
 TTCCATTTTG TTCAGATAAA AAAAACCATG AAATTACTGT GTTTGAAATA TTTTCTTATG 2280  
 GTTTGTAATA TTTCTGTAAA TTTATTGTGA TATTTTAAGG TTTTCCCCC TTTATTTTCC 2340  
 GTAGTTGTAT TTTAAAAGAT TCGGCTCTGT ATTATTTGAA TCAGTCTGCC GAGAATCCAT 2400  
 10 GTATATATTT GAACTAATAT CATCCTTATA ACAGGTACAT TTTCAACTTA AGTTTCTACT 2460  
 CCATTATGCA CAGTTTGAGA TAAATAAATT TTTGAAATAT GGACACTGAA AAAAAAAAAA 2520  
 AAAAAACAA AAAAAAACA CAAAAACAA AAACAGAAAA AACAAAAAA AAAACAAAAC 2580  
 CACAACACAA AAACAAAAA AAAAAAAGA AACAAACACA CAACACAACA CAACACAAAA 2640  
 CCACAACACA AACACAACA CACAGAGGG

Seq ID NO: 4 Protein sequence:  
 Protein Accession #: CAA83435.1

1 11 21 31 41 51  
 | | | | |  
 20 MYNMMETELK PPGPQQTSGG GGGNSTAAAA GGNQKNSPDR VKRPMNAFMV WSRGQRRKMA 60  
 QENPKMHNSE ISKRLGAWEK LLSETEKRPF IDEAKRLRAL HMKEHPDYKY RPRRKTKTLM 120  
 KDKYTLPGG LLAGPGNSMA SGVGVGAGLG AGVNQRMDSY AHMNGWSNGS YSMMQDQLGY 180  
 PQHPGLNAHG AAQMOPMHRV DVSALQYNM TSSQTYMNGS PTYSMSYSQQ GTPGMALGSM 240  
 25 GSVVKSEASS SPPVVTSSSH SRAPCQAGDL RDMISMYLPG AEVPEPAAPS RLHMSQHYQS 300  
 GPVPGTAING TLPLSHM

Seq ID NO: 5 DNA sequence  
 Nucleic Acid Accession #: U91618  
 Coding sequence: 29-541

1 11 21 31 41 51  
 | | | | |  
 30 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60  
 CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTCAAGAG AGGAAATGAA 120  
 35 AGCATTAGAA GCAGATTTCT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180  
 TCCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC 240  
 AGCTGAGGAA ACAGGAGAAG TTCATGAAGA GGAGCTTGTT GCAAGAAGGA AACTTCCTAC 300  
 TGCTTTAGAT GGCTTTAGCT TGGAAAGCAAT GTTGACAATA TACCAGCTCC ACAAATCTG 360  
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA 420  
 40 TGACAAAAAT GAAAGGAAG AAGTATATAA GAGAAAAATT CTTATATATC TGAAACGGCA 480  
 GCTGTATGAG AATAAACCCA GAAGACCCTA CATACTCAAA AGAGATTCTT ACTATTACTG 540  
 AGAGAATAAA TCATTTATTT ACATGTGATT GTGATTCATC ATCCCTTAAT TAAATATCAA 600  
 ATTATATTTG TGTGAAAATG TGACAAACAC ACTTATCTGT CTCTTCTACA ATTGTGGTTT 660  
 45 ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAA TAAATCTAAA 720  
 TCTTCAAAAA AAAAAAAAAA AATGGGGGCC GCAATT

Seq ID NO: 6 Protein sequence:  
 Protein Accession #: AAB50564

1 11 21 31 41 51  
 | | | | |  
 55 MMAGMKIQLV CMLLLAFSSW SLCSDSEEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60  
 VCSLVNNLNS PAEETGEVHE EELVARRKLP TALDGFSLFA MLTIYQLHKI CHSRAFQHWE 120  
 LIQEDILDTG NDKNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSYYY

Seq ID NO: 7 DNA sequence  
 Nucleic Acid Accession #: NM\_006536.2  
 Coding sequence: 109-2940

1 11 21 31 41 51  
 | | | | |  
 65 ACCTAAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60  
 ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120  
 AGCATTGCAG GTCCTATTTG CAACCTGAAG TTTGTGACTC TCCTGGTTGC CTTAAGTTCA 180  
 GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAGACA ATGGGTATAA TGGATTGCTC 240  
 ATTGCAATTA ATCCTCAGGT ACCTGAGAAT CAGAACCCTA TCTCAAACAT TAAGGAAATG 300  
 70 ATAACTGAAG CTTCAATTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTGAGAAAT 360  
 ATAAAGATTT TAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAAGAA 420  
 TCATATGAAA AGGCAAATGT CATAGTGACT GACTGGTATG GGGCACATGG AGATGATCCA 480  
 TACACCTTAC AATACAGAGG GTGTGGAAAA GAGGGAAAAA ACATTCATTT CACACCTAAT 540  
 TTCCTACTGA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCCGAGT GTTTGTCCAT 600  
 75 GAATGGGCCC ACCTCCGTTG GGGTGTGTTT GATGAGTATA ACAATGACAA ACCTTCTTAC 660  
 ATAAATGGGC AAAATCAAAAT TAAAGTGACA AGGTGTTTCT CTGACATCAC AGGCATTTTT 720  
 GTGTGTGAAA AAGGTCCTTG CCCCCAAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780  
 GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840  
 AGTTTATCTT CTGTGTTTGA ATTTTGTAAAT GCAAGTACCC ACAACCAAGA AGCACCAAAC 900  
 CTACAGAACC AGATGTGCAG CCTCAGAAGT GCATGGGATG TAATCACAGA CTCTGCTGAC 960  
 80 TTTCACCACA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCCTCCAC ATTCTCGCTT 1020  
 GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTGCTGGATG TGTCCAGCAA GATGGCAGAG 1080  
 GCTGACAGAC TCCTTCAACT ACAACAAGCC GCAGAATTTT ATTTGATGCA GATTGTTGAA 1140  
 ATTCTACCT TCGTGGGCAT TGCCAGTTTC GACAGCAAGG GAGAGATCAG AGCCCAGCTA 1200  
 CACCAAAATTA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260  
 85 TCAGCTAAAA CAGACATCAG CATTTGTTCA GGGCTTAAGA AAGGATTTGA GGTGGTTGAA 1320  
 AAACCTGAATG GAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG 1380  
 CTTCTTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCATTGCC 1440



5  
10  
15  
20  
25  
30  
35  
40

CTGGGTTTCAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500  
TTCTTTGTTT CAGATATATC AAATCCAAT AGCATGATTG ATGCTTTCAG TAGAATTTC 1560  
TCTGGAACCTG GAGACATTTT CCAGCAACAT ATTCAGCTTG AAAGTACAGG TGAAAATGTC 1620  
AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680  
ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCCTGAGA TTATATTATT TGATCCTGAT 1740  
GGACGAAAAT ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800  
TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCCTGAACAA TACCCATCAT 1860  
TCTCTGCAAG CCCTGAAAGT GACAGTGACC TCTCGCGCCT CCAACTCAGC TGTGCCCCCA 1920  
GCCACTGTGG AAGCCTTTGT GGAAAGAGAC AGCCTCCATT TTCCTCATCC TGTGATGATT 1980  
TATGCCAATG TGAAACAGGG ATTTTATCCC ATTCTTAATG CCACTGTCAC TGCCACAGTT 2040  
GAGCCAGAGA CTGGAGATCC TGTTACGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100  
GTTATAAAAA ATGATGGAAT TTAATCGAGG TATTTTCTT CTTTGCTGTC AAATGGTAGA 2160  
TATAGCTTGA AAGTGCATGT CAATCACTCT CCCAGCATAA GCACCCCAAC CCACTCTATT 2220  
CCAGGGAGTC ATGCTATGTA TGTACCAGGT TACACAGCAA ACGGTAATAT TCAGATGAAT 2280  
GTCCAAGGA AATCAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340  
AGCTCAGGAG GCTCCTTTTC AGTGCTGGGA GTTCCAGCTG GCCCCACCC TGATGTGTTT 2400  
CCACCATGCA AAATTATTGA CCTGGAAGCT GTAAAAGTAG AAGAGGAATT GACCCTATCT 2460  
TGGACAGCAC CTGGAGAAGA CTTTGATCAG GGCCAGGCTA CAAGCTATGA AATAAGAATG 2520  
AGTAAAAGTC TACAGAATAT CCAAGATGAC TTTAACAATG CTATTTTAGT AAATACATCA 2580  
AAGCGAAATC CTCAGCAAGC TGGCATCAGG GAGATATTTA CGTTCTCACC CCAGATTTC 2640  
ACGAATGGAC CTGAACATCA GCCAAATGGA GAAACACATG AAAGCCACAG AATTATGTT 2700  
GCAATACGAG CAATGGATAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCAGGCG 2760  
CCTCTGTTTA TTCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820  
GGAQTTTTAA CAGCAATGGG TTTGATAGGA ATCATTTGCC TTATTATAGT TGTGACACAT 2880  
CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATATATA 2940  
ATAAATATCC AAAGTGTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAAA 3000  
CATACTAACA AAGTCAAAT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060  
ATACAGATAA GATTTTACA TGGTAGATCA ACAATTCTTT TTGGGGGTAG ATTAGAAAAC 3120  
CCTTACACTT TGGCTATGAA CAAATAATAA AAATTATTCT TTAAGTAAT GTCTTTAAAG 3180  
GCAAAGGGAA GGGTAAAGTC GGACCAGTGT CAAGGAAAGT TTGTTTTATT GAGGTGGAAA 3240  
AATAGCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA 3300  
TCATTTAGTT ACTTTGATTA ATTTTCTTT TCTCCTTATC TGTGCAGTAC AGGTTGCTTG 3360  
TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAAG CTCTTACCT 3420  
CTTGCTATTT TGTATATAT ATTTCCCTT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480  
TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540  
TTTATGACAA AGGTCTATTG AATTTATTTG TNTGTAAAGT TCTACTCCCA TCAAAGCAGC 3600  
TTTCTAAGTT TATTGCCTTG GGTATTATAG GAATGATAGT TATAGCCCN TATAATGCCT 3660  
TACCTAGGAA A

Seq ID NO: 8 Protein sequence:  
Protein Accession #: NP\_006527.1

45  
50  
55  
60

1 11 21 31 41 51  
| | | | |  
MTQRSIAGPI CNLKFTLLV ALSSELPFLG AGVQLQDNGY NGLLIAINPQ VPENQNLISN 60  
IKEMITEASF YLFNATKRRV FFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAG 120  
GDDPYTLQYR GCGKEGKYIH FTPNFLNDN LTAGYGSRRG VVHEWAHLR WGVFDEYNND 180  
KPFYINGQNO IKVTRCSSDI TGIFVCEKGP CPQENCIISK LFKEGCTFIY NSTQNATASI 240  
MFMQSLSSVV EFCNASTHNQ EAPNLQNMCM SLRSAWDVIT DSADFHHSPF MNGTELPPPP 300  
TFSLVQAGDK VVCLVLDVSS KMAEADRLLO LQQAEEFYLM QIVEIHTFVG IASFDKGEI 360  
RAQLHQINSN DDRKLLVSYL PTTVSAKTDI SICSGLKKGF EVVEKLNKGA YGSVMILVTS 420  
GDDKLLGNCL PTVLSSSGSTI HSIALGSSAA PNLEELSRLT GGLKFFVDPDI SNSNSMIDAF 480  
SRISSTGDI FQVHIQLEST GENVKPHQL KNTVTVDNTV GNDTMFLVTW QASGPPEIIL 540  
FDPDGRKYIT NNFITNLFR TASLWIPGTA KPGHWYTLN NTHHSLQALK VTVTSRASNS 600  
AVPPATVEAF VERDSLHFPV PVMYANVKQ GFYPILNATV TATVEPETGD PVTLRLLDDG 660  
AGADVIKNDG IYSRYFFSFA ANGRYSKLVH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720  
IQMNAFRKSV GRNEEERKWG FSRVSSGGSF SVLGVPAGPH PDVFPCKII DLEAVKVEEE 780  
LTLSTWAPGE DFDQGGQTSY EIRMSKSLQN IQDDFNAIL VNTSKRNPQQ AGIREIFTFS 840  
PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900  
LILKGVLTAM GLIGIICLII VVTHHTLSRK KRADKKENG T KLL

Seq ID NO: 9 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 336-632

70  
75  
80  
85

1 11 21 31 41 51  
| | | | |  
CTCCCCCTCAC CCCGGTCCAG GATGCCCAGT CCCACGACA CCTCCCCTT CCCACTGTGG 60  
CCTGGGTGGG CTCAGGGGCT GCCCTTGACC TGGCCTAGAG CCCTCCCCCA GCTGGTGGTG 120  
GAGCTGGCAC TCTCTGGGAG GGAGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180  
CAGGGTTTGG TGGGATCAGG TTGAGGCAGG TTTGGTTTCC TTAATAATGCC AAGTTGGGG 240  
CCAGTGGGGC CCACATATAA ATCCTCACCC TGGGAGCCTG GCTGCCTTGC TCTCCTTCCT 300  
GGGTCTGTCT CTGCCACCTG GTCTGCCACA GATCCATGAT GTGCAGTTCT CTGGAGCAGG 360  
CGCTGGCTGT GCTGGTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420  
AGCTGAGTAA GGGGGAATG AAGGAACCTC TGCACAAGGA GCTGCCAGC TTTGTGGGG 480  
AGAAAGTGGA TGAGGAGGGG CTGAAGAAGC TGATGGGCAG CCTGGATGAG AACAGTGACC 540  
AGCAGGTGGA CTTCCAGGAG TATGCTGTTT TCCTGGCACT CATCACTGTC ATGTGCAATG 600  
ACTTCTTCCA GGGCTGCCCA GACCGACCCT GAAGCAGAAC TCTTGACTTC CTGCCATGGA 660  
TCTCTTGGGC CCAGGACTGT TGATGCCTTT GAGTTTGTG TTTCAATAAAC TTTTGTGTC 720  
TGTTGATAAT ATTTTAATTG CTCAGTGATG TTCCATAACC CGGCTGGCTC AGCTGGAGTG 780  
CTGGGAGATG AGGGCCTCCT GGATCCTGCT CCCTTCTGGG CTCTGACTCT CCTGGAAATC 840  
TCTCCAAGGC CAGAGCTATG CTTTAGGTCT CAATTTTGA ATTTCAAACA CCAGCAAAAA 900  
ATTGGAAATC GAGATAGGTT GCTGACTTTT ATTTTGTCAA ATAAAGATAT TAAAAAGGC 960  
AAATACCA

Seq ID NO: 10 Protein sequence:



Protein Accession #: NP\_005969.1

5                   1                   11                   21                   31                   41                   51  
|                   |                   |                   |                   |                   |  
MMCSSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGEK VDEEGLKKLM 60  
GSLDENSDQQ VDFQEYAVFL ALITVMCNDF FQGCPDRP

10 Seq ID NO: 11 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 336-626

15                   1                   11                   21                   31                   41                   51  
|                   |                   |                   |                   |                   |  
CTCCCCCTCAC CCCGGTCCAG GATGCCCAGT CCCACGACA CCTCCCACTT CCCACTGTGG 60  
CCTGGGTGGG CTCAGGGGCT GCCCTTGACC TGGCCTAGAG CCCTCCCCCA GCTGGTGGTG 120  
GAGCTGGCAC TCTCTGGGAG GGAGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180  
CAGGGTTTGG TGGGATCAGG TTGAGGCAGG TTTGGTTTCC TTAAAATGCC AAGTTGGGGG 240  
CCAGTGGGGC CCACATATAA ATCCTCACCC TGGGAGCCTG GCTGCCTTGC TCTCCTTCCT 300  
GGGTCTGTCT CTGCCACCTG GTCTGCCACA GATCCATGAT GTGCAGTTCT CTGGAGCAGG 360  
CGCTGGCTGT GCTGGTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420  
AGCTGAGTAA GGGGGAAATG AAGGAACTTC TGCACAAGGA GCTGCCAGC TTTGTGGGGC 480  
ATCCAGAGA ACCATGTGCT GTGAGGGCCT TCCGAGTCCA TCTGTTTAAT CCTGTTCATTG 540  
GAGACTTGAG AAACCAGAGC CCAGAAGGGA AAAGTGATTG TCCCAAGATC ACACAGCACT 600  
25 GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC AGCCTGGATG AGAACAGTGA 660  
CCAGCAGGTG GACTTCCAGG AGTATGCTGT TTTCTGGCA CTCATCACTG TCATGTGCAA 720  
TGACTTCTTC CAGGGCTGCC CAGACCGACC CTGAAGCAGA ACTCTTGACT TCCTGCCATG 780  
GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG TATTCAATAA ACTTTTTTTG 840  
30 TCTGTTGATA ATATTTTAAT TGCTCAGTGA TGTTCATAA CCCGGCTGGC TCAGCTGGAG 900  
TGCTGGGAGA TGAGGGCCTC CTGGATCCTG CTCCCTTCTG GGCTCTGACT CTCTGGAAA 960  
TCTCTCAAG GCCAGAGCTA TGCTTTAGGT CTCAATTTTG GAATTTCAA CACCAGCAA 1020  
AAATTGAAA TCGAGATAGG TTGCTGACTT TTATTTTGTC AAATAAAGAT ATTAAAAAAG 1080  
GCAAATACCA

35 Seq ID NO: 12 Protein sequence:  
Protein Accession #: Eos sequence

40                   1                   11                   21                   31                   41                   51  
|                   |                   |                   |                   |                   |  
MMCSSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGHS REPCAVRAFR 60  
VHLFNPVIGD LRNQSPGKKS DCPKITQHR KWMRRG

45 Seq ID NO: 13 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 58-354

50                   1                   11                   21                   31                   41                   51  
|                   |                   |                   |                   |                   |  
GTGAGCTCAC CATGTGGGGG TGAGGCTGAG AGAAAACAAG TACACAGCCA CAGATCCATG 60  
ATGTGCAGTT CTCTGGAGCA GCGCTGGCT GTGCTGGTCA CTACCTTCCA CAAGTACTCC 120  
TGCCAAGAGG GCGACAAGTT CAAGCTGAGT AAGGGGAAA TGAAGGAAT TCTGCACAAG 180  
GAGCTGCCCA GCTTTGTGGG GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC 240  
55 AGCCTGGATG AGAACAGTGA CCAGCAGGTG GACTTCCAGG AGTATGCTGT TTTCTGGCA 300  
CTCATCACTG TCATGTGCAA TGACTTCTTC CAGGGCTGCC CAGACCGACC CTGAAGCAGA 360  
ACTCTTGACT TCCTGCCATG GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG 420  
TATTCAATAA ACTTTTTTTG TCTGTTGATA ATATTTTAAT TGCTCAGTGA TGTTCATAA 480  
CCCGGCTGGC TCAGCTGGAG TGCTGGGAGA TGAGGGCCTC CTGGATCCTG CTCCCTTCTG 540  
60 GGCTCTGACT CTCCTGAAA TCTCTCCAAG GCCAGAGCTA TGCTTTAGGT CTCAATTTTG 600  
GAATTTCAA CACCAGCAA AAATTGAAA TCGAGATAGG TTGCTGACTT TTATTTTGTC 660  
AAATAAAGAT ATTAAAAAAG GCAAATACCA

65 Seq ID NO: 14 Protein sequence:  
Protein Accession #: NP\_005969.1

70                   1                   11                   21                   31                   41                   51  
|                   |                   |                   |                   |                   |  
MMCSSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGEK VDEEGLKKLM 60  
GSLDENSDQQ VDFQEYAVFL ALITVMCNDF FQGCPDRP

75 Seq ID NO: 15 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 62-358

80                   1                   11                   21                   31                   41                   51  
|                   |                   |                   |                   |                   |  
GGAGGGTGTG CCGCTGAGTC ACTGCCTGGG CATCTGGGCC TGGAACCTCG GCCACAGATC 60  
CATGATGTGC AGTTCTCTGG AGCAGGCGCT GGCTGTGCTG GTCACTACCT TCCACAAGTA 120  
CTCCTGCCAA GAGGGCGACA AGTTCAAGCT GAGTAAGGGG GAAATGAAGG AACTTCTGCA 180  
85 CAAGGAGCTG CCCAGCTTTG TGGGGGAGAA AGTGGATGAG GAGGGGCTGA AGAAGCTGAT 240  
GGGCAGCCTG GATGAGAACA GTGACCAGCA GGTGGACTTC CAGGAGTATG CTGTTTTCTT 300  
GGCACTCATC ACTGTCATGT GCAATGACTT CTTCAGGGC TGCCCAGACC GACCCTGAAG 360  
CAGAACTCTT GACTTCTTGC CATGGATCTG TTGGGCCAG GACTGTTGAT GCCTTTGAGT 420  
TTTGTAATTA ATAACTTTT TTTGTCTGTT GATAATATT TAATTGCTCA GTGATGTTCC 480  
ATAACCCGCG TGGCTCAGCT GGAGTGCTGG GAGATGAGGG CCTCTGGAT CCTGCTCCCT 540  
TCTGGGCTCT GACTCTCTCT GAAATCTCTC CAAGGCCAGA GCTATGCTTT AGGTCTCAAT 600  
TTTGGAATTT CAAACACCAG CAAAAAATTG GAAATCGAGA TAGGTTGCTG ACTTTTATTT 660



TGTCAAATAA AGATATTAA AAAGGCAAAT ACCA

Seq ID NO: 16 Protein sequence:  
Protein Accession #: NP\_005969.1

1	11	21	31	41	51	
MMCSSLEQAL	AVLVTTFHKY	SCQEGDKFKL	SKGEMKELLH	KELPSFVGEK	VDEEGLKKLM	60
GSLDENSDDQ	VDFQEYAVFL	ALITVMCNDF	FQGCPRDP			

Seq ID NO: 17 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 939-2372

1	11	21	31	41	51	
AAGACGGATT	CTCAGACAAG	GCTTGCAAAT	CCCCCGCAGC	CATCATTTAA	CTGCACCCGC	60
AGAATAGTTA	CGGTTTGTCA	CCCGACCCTC	CCGGATCGCC	TAATTTGTCC	CTAGTGAGAC	120
CCCGAGGCTC	TGCCCCGCGC	TGGCTTCTTC	GTAGCTGGAT	GCATATCGTG	CTCCGGGCAG	180
CGCGGGCGCA	GGGCACGCGT	TCGCGCACAC	CCTAGCACAC	ATGAACACGC	GCAAGAGCTG	240
AACCAAGCAC	GGTTTCCATT	TCAAAAAGGG	AGACAGCCTC	TACCGCGATT	GTAAGAGAGA	300
CTGTGGTGTG	AATTAGGGAC	CGGGAGGCGT	CGAACGGAGG	AACGGTTCAT	CTTAGAGACT	360
AATTTTCTGG	AGTTTCTGCC	CCTGCTCTGC	GTACAGCCCTC	ACGTCACTTC	GCCAGCAGTA	420
GCAGAGGCGG	CGGCGGCGGC	TCCCAGGAAT	GGGTTGGAGC	AGGAGCCTCG	CTGGCTGCTT	480
CGCTCGCGCT	CTACGCGCTC	AGTCCCCGGC	GGTAGCAGGA	GCCTGGACCC	AGGCGCCGCC	540
GGCGGGCGTG	AGGCGCCGGA	GCCCCGCCCTC	GAGGTGCATA	CCGGACCCCC	ATTGCGATCT	600
AACAAGGAAT	CTGCGCCCCA	GAGAGTCCCG	GGAGCGCCGC	CGGTCCGGTG	CCGGCGCGCC	660
GGGCCATGCA	GCGACGGCCG	CCGCGGAGCT	CCGAGCAGCG	GTAGCGCCCC	CCTGTAAAGC	720
GGTTCCGCTAT	GCCGGGGCCA	CTGTGAACCC	TGCCGCCTGC	CGGAACACTC	TTCGCTCCGG	780
ACCAGCTTCAG	CCTCTGATAA	GCTGGACTCG	GCACGCCCGC	AACAAGCACC	GAGGAGTTAA	840
GAGAGCCGCA	AGCGCAGGGA	AGGCCTCCCC	GCACGGGTGG	GGGAAAGCGG	CCGGTGCAGC	900
GCGGGGACAG	GCACTCGGGC	TGGCACTGGC	TGCTAGGGAT	GTCTCCTGG	ATAAGGTGGC	960
ATGGACCCGC	CATGGCGCGG	CTCTGGGGCT	TCTGCTGGCT	GGTTGTGGGC	TTCTGGAGGG	1020
CCGCTTTCGC	CTGTCCCACG	TCCTGCAAAT	GCAGTGCCCTC	TCCGATCTGG	TGCAGCGACC	1080
CTTCTCCTGG	CATCGTGGCA	TTTCCGAGAT	TGGAGCCTAA	CAGTGTAGAT	CCTGAGAACA	1140
TCACCGAAAT	TTTCATCGCA	AACCGAATAA	GGTTAGAAAT	CATCAACGAA	GATGATGTTG	1200
AAGCTTATGT	GGGACTGAGA	AATCTGACAA	TTGTGATTC	TGGATTAAAA	TTTGTGGCTC	1260
ATAAAGCATT	TCTGAAAAAC	AGCAACCTGC	AGCACATCAA	TTTTACCCGA	AACAACTGA	1320
CGAGTTTGTG	TAGGAAACAT	TTCCGTCACC	TTGACTTGTG	TGAACTGATC	CTGGTGGGCA	1380
ATCCATTTAC	ATGCTCCTGT	GACATTATGT	GGATCAAGAC	TCTCCAAGAG	GCTAAATCCA	1440
GTCCAGACAC	TCAGGATTGT	TACTGCCTGA	ATGAAAGCAG	CAAGAATATT	CCCCTGGCAA	1500
ACCTGCAGAT	ACCCAATTGT	GGTTTGCCAT	GTGCAATCT	GGCCGCACCT	AACCTCACTG	1560
TGGAGGAAGG	AAAGTCTATC	ACATTATCCT	GTAGTTGGGC	AGGTGATCCG	GTTCCTAATA	1620
TGTATTGGGA	TGTTGGTAAC	CTGGTTTCCA	AACATATGAA	TGAAACAAGC	CACACACAGG	1680
GCTCCTTAAG	GATAACTAAC	ATTTTCATCCG	ATGACAGTGG	GAAGCAGATC	TCTTGTGTGG	1740
CGGAAAATCT	TGTAGGAGAA	GATCAAGATT	CTGTCAACCT	CACTGTGCAT	TTTGCACCAA	1800
CTATCACATT	TCTCGAATCT	CCAACCTCAG	ACCACCACTG	GTGCATTCCA	TTCCTGTGA	1860
AAGGCAACCC	CAAAACAGCG	CTTCAGTGGT	TCTATAACGG	GGCAATATTG	AATGAGTCCA	1920
AATACATCTG	TACTTAAATA	CATGTTACCA	ATCACACGGA	GTACCACGGC	TGCCTCCAGC	1980
TGGATAATCC	CACTCACATG	AACAATGGGG	ACTACACTCT	AATAGCCAAG	AATGAGTATG	2040
GGAAGGATGA	GAAACAGATT	TCTGCTCACT	TCATGGGCTG	GCCTGGAATT	GACGATGGTG	2100
CAAACCCAAA	TTATCCTGAT	GTAATTTATG	AAGATTATGG	AACTGCAGCG	AATGACATCG	2160
GGGACACCC	GAACAGAAAT	AATGAAATCC	CTTCCACAGA	CGTCACTGAT	AAAACCGGTC	2220
GGGAACATCT	CTCGGTCTAT	GCTGTGGTGG	TGATTGCGTC	TGTGGTGGGA	TTTGGCCTTT	2280
TGGTAATGCT	GTTTCTGCTT	AAGTTGGCAA	GACACTCCAA	GTTTGGCATG	AAAGGTTTGT	2340
TTTTGTTTCA	TAAGATCCCA	CTGGATGGGT	AGCTGAAATA	AAGGAAAAGA	CAGAGAAAGG	2400
GGCTGTGGTG	CTTGTGGTGT	GATGCTGCCA	TGTAAGCTGG	ACTCCTGGGA	CTGCTGTTGG	2460
CTTATCCCGG	GAAGTGCTGC	TTATCTGGGG	TTTTCTGGTA	GATGTGGGCG	GTGTTTGGAG	2520
GCTGTACTAT	ATGAAGCCTG	CATATACTGT	GAGCTGTGAT	TGGGGAACAC	CAATGCAGAG	2580
GTAACCTCTCA	GGCAGCTAAG	CAGCACCTCA	AGAAAACATG	TTAAATTAAT	GCTTCTCTTC	2640
TTACAGTAGT	TCAAATACAA	AACTGAAATG	AAATCCCAT	GGATTGTACT	TCTCTTCTGA	2700
AAAGTGTGCT	TTTTGACCTT	ACTGGACATT	TATTGACTTA	ATTGCTTCTG	TTTATTAAAA	2760
TTGACCTGCA	AAGTTAAAAA	AAAATTAAG	TTGAGAACAG	GTATAAGTGC	ACACTGAATA	2820
GTCTAATCTA	CATGTAACAC	ATATTTTAGT	GTGATTTTCT	ATACTCTAAT	CAGCACTGAA	2880
TTGAGAGGGT	TTGACTTTTT	CATCTATAAC	ACAGTGACTA	AAAGAGTTAA	GGGTATATAT	2940
ACCATCACTT	TGGGACTTGG	TAGTATTATT	AAAAGGTTAT	TTCTTCACT	GTCAATAAAA	3000
GTCCAAATGT	TTAGCTTAGG	TCTGAGAGTC	AAACAATGTT	AAGGATTGTC	TTAAAGTTCC	3060
TTAGCCAGCA	AAACAAAACA	AAACAAAACA	AACAAATGAA	AAACGTTTAA	AAAGAAGAAG	3120
AAGAAAAAAA	ACAAGAACAA	GCAGCAACAG	CTGTTTGTGT	GGGGCTATAG	ATTTAAGTTA	3180
GGCATACTCA	ATTTTCAAGT	AACTAAGAGT	GGAATATATG	CATATGGTGA	AATTATAACC	3240
TTGCCCTTTT	TTATTTGCC	TCTGCGATCC	ACCTGCTTTT	TAGAAGTCTG	CCGAGTGAGA	3300
AGGCCACAGT	ATCTCATGCT	GTTTGCATTA	CAGAACTGCA	GCTTTTCTAC	TCTGAAAAGG	3360
CCTGGGAGCA	GAATGGCTGG	CCTGCTGTGG	GCAGGAGAGG	AGATTCTAAG	AAGGATAGTC	3420
CCCCCTACAA	CATGCTGTCA	TACTGCTGGG	TTTTTCATGG	TAGGAAAGCT	TGTCCTGACC	3480
CCAGCAGCAA	AGAGGTGGCA	GGTCGCTAAT	GAATATATGC	TTTATAATGT	CCTTCTTCAT	3540
TGCTGAGAGG	GCAGCCTTAG	AGCTGTGGAT	TTCTGCATCC	CCCCTGAGTC	TGACCCATGG	3600
ACACCTGTTT	CATTCACCTT	AGCATCACAG	TGACCTTTGT	ATGCTCTGTT	CAGTCTGTGT	3660
CAGGCAGTAT	GCTTGTCTTG	AAGAGAGGTT	TGGCTATCCC	CACCCACCC	CACCCACCC	3720
TGTTCCCTTT	TTATCAGGAG	GACTTCAGAG	CCAGGCCTGC	AGCATTTTGT	TTGAAAACAC	3780
AATCAGCTCT	GACAGTTAGA	CATGCACACA	GACGCCATAG	CTGGATTGGA	AACATTGATG	3840
TTTTAAAAAT	TTATTTTTTT	TGGAAATAGT	TGCACAAATG	CTGCAATTTA	GCTTTAAGGT	3900
TCTATAGATT	TTTAACTAGT	CCAACACAGT	CAGAAACATT	GTTTTGAATC	CTCTGTAAAC	3960
CAAGGCATTA	ATCTTAATAA	ACCAGGATCC	ATTTAGGTAC	CACTTGATAT	AAAAAGGATA	4020
TCCATAATGA	ATATTTTATA	CTGCATCCTT	TACATTAGCC	ACTAAATACG	TTATTGCTTG	4080
ATGAAGACCT	FTCACAGAAAT	CCTATGGATT	GCAGCATTTT	ACTTGGCTAC	TTCATACCCA	4140



5  
10

TGCCTTAAAG	AGGGGCAGTT	TCTCAAAAGC	AGAAACATGC	CGCCAGTTCT	CAAGTTTTTC	4200
TCCTAACTCC	ATTTGAATGT	AAGGGCAGCT	GGCCCCCAAT	GTGGGGAGGT	CCGAACATTT	4260
TCTGAATTCC	CATTTTCTTG	TTTCGCGGCTA	AATGACAGTT	TCTGTCATTA	CTTAGATTCC	4320
GATCTTTCCC	AAAGGTGTTG	ATTTACAAAG	AGGCCAGCTA	ATAGCAGAAA	TCATGACCCT	4380
GAAAGAGAGA	TGAAATTCAA	GCTGTGAGCC	AGGCAGGAGC	TCAGTATGGC	AAAGGTTCTT	4440
GAGAATCAGC	CATTTGGTAC	AAAAAAGATT	TTTAAAGCTT	TTATGTTATA	CCATGGAGCC	4500
ATAGAAAGGC	TATGGATTGT	TTAAGAACTA	TTTAAAGTG	TTCCAGACCC	AAAAAGGAAA	4560
AATAAAAAAA	AAGGAATATT	TGTACCCAAC	AGCTAGAAGG	ATTGCAAGGT	AGATTTTGT	4620
TTTAAATGG	AGAGAAGTGG	ACAGATAAGG	CCATTTAATA	TATCAAAGAT	CAGTTGACAT	4680
CTCCTAGGGA	ATGATGAAAA	CAGCAGGCTA	T			

Seq ID NO: 18 Protein sequence:  
Protein Accession #: CAA53571

15  
20  
25

1	11	21	31	41	51	
MSSWIRWHGP	AMARLWGFCE	LVVGFWRAAF	ACPTSCKCSA	SRIWCSDPSP	GIVAFPRLEP	60
NSVDPENITE	IFIANQKRLE	IINEDDVEAY	VGLRNLITVD	SGLKFAVHKA	FLKNSNLQHI	120
NFTRNKLTSL	SRKHFRHLDL	SELILVGNPF	TCSCDIMWIK	TLQEAQSSPD	TQDLYCLNES	180
SKNIPLANLQ	IPNCGLPAN	LAAPNLTVEE	GKSITLSCSV	AGDPVPMYMW	DVGNLVSXHM	240
NETSHTQSL	RITNISSDDS	GKQISCVAEN	LVGEDQDSVN	LTVHFAPTIT	FLESPTSDDH	300
WCTPFTVKG	NPKPALQWYN	GAILNESKYI	CTKIHVNTHT	EYHGCLQLDN	PTHMNNGDYT	360
LIANEYKGD	EKQISAHFMG	WPGIDDGANP	NYPDVIYEDY	GTAANDIGDT	TNRSNEIPST	420
DVTDKTGREH	LSVYAVVVIA	SVVGFCLLVM	LFLKLKLARHS	KFGMKGFVLF	HKIPLDGL	

Seq ID NO: 19 DNA sequence  
Nucleic Acid Accession #: NM\_000228  
Coding sequence: 82-3600

30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
GCTTTCAGGC	GATCTGGAGA	AAGAACGGCA	GAACACACAG	CAAGGAAAGG	TCCTTTCTGG	60
GGATCACCCC	ATTGGCTGAA	GATGAGACCA	TTCTTCCTCT	TGTGTTTTGC	CCTGCCTGGC	120
CTCCTGCATG	CCCAACAAGC	CTGCTCCCGT	GGGGCCTGCT	ATCCACCTGT	TGGGGACCTG	180
CTTGTTGGGA	GGACCCGGTT	TCTCCGAGCT	TCACTACCT	GTGGACTGAC	CAAGCCTGAG	240
ACCTACTGCA	CCCAGTATGG	CGAGTGGCAG	ATGAAATGCT	GCAAGTGTGA	CTCCAGGCAG	300
CCTCACAACT	ACTACAGTCA	CCGAGTAGAG	AATGTGGCTT	CATCCTCCGG	CCCCATGCGC	360
TGGTGGCAGT	CCCAGAATGA	TGTGAACCC	GTCTCTCTGC	AGCTGGACCT	GGACAGGAGA	420
TTCCAGCTTC	AAGAAGTCAT	GATGGAGTTC	CAGGGGCCCA	TGCCCGCCGG	CATGCTGATT	480
GAGCGCTCCT	CAGACTTCGG	TAGACCTGG	CGAGTGTACC	AGTACCTGGC	TGCCGACTGC	540
ACCTCCACCT	TCCCTCGGGT	CCGCCAGGGT	CGGCTCAGA	GCTGGCAGGA	TGTTCCGGTG	600
CAGTCCCTGC	CTCAGAGGCC	TAATGCACGC	CTAAATGGGG	GGAAGGTCCA	ACTTAACCTT	660
ATGGATTTAG	TGTCTGGGAT	TCCAGCAACT	CAAAGTCAA	AAATTCAAGA	GGTGGGGGAG	720
ATCACAACT	TGAGAGTCAA	TTTACCAGG	CTGGCCCTG	TGCCCCAAAG	GGGCTACCAC	780
CCTCCCAGCG	CCTACTATGC	TGTGTCCCAG	CTCCGTCTGC	AGGGGAGCTG	CTTCTGTCA	840
GGCCATGCTG	ATCGCTGCGC	ACCCAAGCCT	GGGGCCTCTG	CAGGCCCCCT	CACCGCTGTG	900
CAGTCCACG	ATGCTGTGT	CTGCCAGCAC	AACACTGCCG	GCCCAAATTG	TGAGCGCTGT	960
GCACCCCTCT	ACAACAACCG	GCCCTGGAGA	CCGGCGGAGG	GCCAGGACGC	CCATGAATGC	1020
CAAAGGTGCG	ACTGCAATGG	GCACTCAGAG	ACATGTCACT	TTGACCCCGC	TGTGTTTGCC	1080
GCCAGCCAGG	GGGCATATGG	AGGTGTGTGT	GACAATTGCC	GGGACCACAC	CGAAGGCAAG	1140
AACTGTGAGC	GGTGTGAGCT	GCACTATTTT	CGGAACCGGC	GCCCGGGAGC	TTCCATTGAG	1200
GAGACCTGCA	TCTCCTGCGA	GTGTGATCCG	GATGGGGCAG	TGCCAGGGGC	TCCCTGTGAC	1260
CCAGTGACCG	GGCAGTGTGT	GTGCAAGGAG	CATGTGCAGG	GAGAGCGCTG	TGACCTATGC	1320
AAGCCGGGCT	TCACTGGACT	CACCTACGCC	AACCCGAGG	GCTGCCACCG	CTGTGACTGC	1380
AACATCCTGG	GGTCCCGGAG	GGACATGCCG	TGTGACGAGG	AGAGTGGGCG	CTGCCTTTGT	1440
CTGCCCAACG	TGGTGGGTCC	CAAATGTGAC	CAGTGTGCTC	CCTACCACTG	GAAGCTGGCC	1500
AGTGGCCAGG	GCTGTGAACC	GTGTGCCTGC	GACCCGCACA	ACTCCCTCA	GCCCACAGTG	1560
CAACCAGTTC	ACAGGGCAGT	GCCCTGTCCG	GAAGGCTTTG	GTGGCCTGAT	GTGCAGCGCT	1620
GCAGCCATCC	GCCAGTGTCC	AGACCGGACC	TATGGAGACG	TGGCCACAGG	ATGCCGAGCC	1680
TGTGACTGTG	ATTTCGGGGG	AACAGAGGGC	CCGGCTGCG	ACAAGGCATC	AGGCCGCTGC	1740
CTCTGCCGCC	CTGGCTTGAC	CGGGCCCCGC	TGTGACCAGT	GCCAGCGAGG	CTACTGCAAT	1800
CGCTACCCGG	TGTGCGTGGC	CTGCCACCCT	TGCTTCCAGA	CCTATGATGC	GGACCTCCGG	1860
GAGCAGGCCC	TGCGCTTTGG	TAGACTCCGC	AATGCCACCG	CCAGCCTGTG	GTGAGGGCCT	1920
GGGCTGGAGG	ACCGTGGCCT	GGCCTCCCGG	ATCCTAGATG	CAAAGAGTAA	GATTGAGCAG	1980
ATCCGAGCAG	TTCTCAGCAG	CCCCGACCTG	ACAGAGCAGG	AGGTGGCTCA	GGTGGCCAGT	2040
GCCATCCTCT	CCCTCAGGCG	AACCTCTCAG	GGCCTGCAGC	TGGATCTGCC	CCTGGAGGAG	2100
GAGACGTTGT	CCCTTCCGAG	AGACCTGGAG	AGTCTTGACA	GAAGCTTCAA	TGGTCTCCTT	2160
ACTATGTATC	AGAGGAAGAG	GGAGCAGTTT	GAAAAAATAA	GCAGTGCTGA	TCCTTCAGGA	2220
GCCTTCCGGA	TGCTGAGCAC	AGCCTACGAG	CAGTCAGCCC	AGGCTGCTCA	GCAGGTCTCC	2280
GACAGCTCGC	GCCTTTTGGA	CCAGCTCAGG	GACAGCCGGA	GAGAGGCAGA	GAGGCTGGTG	2340
CGGCAGGCGG	GAGGAGGAGG	AGGCACCGGC	AGCCCCAAGC	TTGTGGCCCT	GAGGCTGGAG	2400
ATGTCTTCGT	TGCCCTGACCT	GACACCCACC	TTCAACAAGC	TCTGTGGCAA	CTCCAGGCAG	2460
ATGGCTTGCA	CCCCAATATC	ATGCCCTGGT	GAGCTATGTC	CCCAAGACAA	TGGCACAGCC	2520
TGTGGCTCCC	GCTGCAGGGG	TGTCCTTCCC	AGGGCCGGTG	GGGCCTTCTT	GATGGCGGGG	2580
CAGGTGGCTG	AGCAGCTGCG	GGGCTTCAAT	GCCCAGCTCC	AGCGGACCAG	GCAGATGATT	2640
AGGGCAGCCG	AGGAATCTGC	CTCACAGATT	CAATCCAGTG	CCCAGCGCTT	GGAGACCCAG	2700
GTGAGCGCCA	GCCGCTCCCA	GATGGAGGAA	GATGTCAGAC	GCACACGGCT	CCTAATCCAG	2760
CAGGTCCGGG	ACTTCCTAAC	AGACCCGAG	ACTGATGCAG	CCACTATCCA	GGAGGTGAGC	2820
GAGGCCGTGC	TGGCCCTGTG	GCTGCCCCAC	ACTCAGCTA	CTGTTCTGCA	GAAGATGAAT	2880
GAGATCCAGG	CCATTGCAGC	CAGGCTCCCC	AACGTGGACT	TGGTGTCTGC	CCAGACCAAG	2940
CAGGACATTG	CGCGTGCCCG	CCGGTTGCAG	GCTGAGGCTG	AGGAAGCCAG	GAGCCGAGCC	3000
CATGCAGTGG	AGGGCCAGGT	GGAAGATGTG	GTTGGGAACC	TGCGGCAGGG	GACAGTGGCA	3060
CTGCAGGAAG	CTCAGGACAC	CATGCAAGGC	ACCAGCCGCT	CCCTTCGGCT	TATCCAGGAC	3120
AGGGTTGCTG	AGGTTCAGCA	GGTACTGCGG	CCAGCAGAAA	AGCTGGTGAC	AAGCATGACC	3180
AAGCAGCTGG	GTGACTTCTG	GACACGGATG	GAGGAGCTCC	GCCACCAAGC	CCGGCAGCAG	3240
GGGGCAGAGG	CAGTCCAGGC	CCAGCAGCTT	GCGGAAGGTG	CCAGCGAGCA	GGCATTGAGT	3300
GCCCAAGAGG	GATTTGAGAG	AATAAAACAA	AAGTATGCTG	AGTTGAAGGA	CCGGTTGGGT	3360



CAGAGTTCCA TGCTGGGTGA GCAGGGTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG 3420  
 GAGCTGTTTG GGGAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480  
 CTGCGGGGCA GCCAGGCCAT CATGCTGCGC TCGGCGGACC TGACAGGACT GGAGAAGCGT 3540  
 GTGGAGCAGA TCCGTGACCA CATCAATGGG CGCGTGCTCT ACTATGCCAC CTGCAAGTGA 3600  
 5 TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCCGCC TTTGCTTTTG GTTGGGGGCA 3660  
 GATTGGGTTG GAATGCTTTC CATCTCCAGG AGACTTTCAT GCAGCCTAAA GTACAGCCTG 3720  
 GACCACCCCT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAAT 3780  
 GGGACAGTTA CACTTGACAG ACAAAGATGG TGGAGATTGG CATGCCATTG AAACATAAGAG 3840  
 10 CTCTCAAGTC AAGGAAGCTG GGCTGGGCAG TATCCCCGCG CTTTAGTTCT CCACTGGGGA 3900  
 GGAATCCTGG ACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAAATGAA AAGCCAAATA 3960  
 AAAATCTTTG G

Seq ID NO: 20 Protein sequence:  
 Protein Accession #: NP\_000219

1 11 21 31 41 51  
 | | | | |  
 MRPFLLCFA LPGLLHAQQA CSRGACYPPV GDLLVGRTRF LRASSTCGLT KPETYCTQYG 60  
 EWQMKCKCD SRQPHNYSH RVENVASSSG PMRWWQSQND VNPVSLQLDL DRRFQLQEVN 120  
 20 MEFQGPMPAG MLIERSSDFG KTWRYVQYLA ADCTSTFPRV RQGRPQSWQD VRCQSLPQRP 180  
 NARLNGGKVQ LNLMDLVSGI PATQSQKIQE VGEITNLRVN FTRLAPVPQR GYHPPSAYYA 240  
 VSQRLRQSGC FCHGHADRC PKPGASAGPS TAVQVHDVCV CQHNTAGPNC ERCAPFYNNR 300  
 PWRPAEGQDA HECQRCDNG HSETCHFDPA VFAASQGAYG GVCDNCRDHT EGKNCERCQL 360  
 25 HYFRNRRPGA SIQETCISCE CDPDGAVPGA PCDPVTGQCV CKEHVQGERC DLCKPGFTGL 420  
 TYANPQGCHR CDCNILGSRD DMPDDEESGR CLCLPNVVGK KCDQCAPHYH KLASGQGCPE 480  
 CACDPHNSPQ PTVQPVHRAV PCREGFGGLM CSAAAIRQCP DRTYGDVATG CRACDCDFRG 540  
 TEGPGCDKAS GRCLCRPGLT GPRCDQCQRG YCNRYPCVCA CHPCFQTYDA DLREQALRFG 600  
 RLRNATASLW SGPGLDRGL ASRILDASK IEQIRAVLSS PAVTEQEVQA VASAILSLRR 660  
 30 TLQGLQLDLP LEEETLSLPR DLESIDRSFN GLLTMYQRKR EQFEKISSAD PSGAFRMLST 720  
 AYEQSAQAAQ QVSDSSRLLD QLRDSRREAE RLVRQAGGGG GTGSPKLVAL RLEMSSLPDL 780  
 TPTFNKLCGN SRQMACTPIS CPGELCPQDN GTACGSRERG VLPRAGGAFL MAGQVAEQLR 840  
 GFNAQLQRTQ QMIRAAEESA SQIQSSAQRL ETQVSASRSQ MEEDVRRTRL LIQQVRDFLT 900  
 DPDTDAATIQ EVSEAVLALW LPTDSATVLQ KMNEIQAIQA RLPNVDLVLS QTKQDIARAR 960  
 35 RLQAEAEER SRAHAVEGQV EDVVGNLRQG TVALQEAQDT MQGTSRSLRL IQDRVAEVQQ 1020  
 VLRPAEKLVT SMTKQLGDFW TRMEELRHQA RQQAEGAVQA QQLAEGASEQ ALSAQEGFER 1080  
 IKQKYAELKD RLQSSMLGE QGARIQSVKT EAEELFGETM EMMDRMKDME LELLRGSQAI 1140  
 MLRSADLTGL EKRVEQIRDH INGRVLYYAT CK

Seq ID NO: 21 DNA sequence  
 Nucleic Acid Accession #: NM\_003722  
 Coding sequence: 145-1491

1 11 21 31 41 51  
 | | | | |  
 TCGTTGATAT CAAAGACAGT TGAAGGAAAT GAATTTTGAA ACTTCACGGT GTGCCACCCT 60  
 ACAGTACTGC CCTGACCCTT ACATCCAGCG TTTCGTAGAA ACCCAGCTCA TTTCTCTTGG 120  
 AAAGAAAGTT ATTACCGATC CACCATGTCC CAGAGCACAC AGACAAATGA ATTCCTCAGT 180  
 CCAGAGGTTT TCCAGCATAT CTGGGATTTT CTGGAACAGC CTATATGTTT AGTTCAGCCC 240  
 50 ATTGACTTGA ACTTTGTGGA TGAACCATCA GAAGATGGTG CGACAAACAA GATTGAGATT 300  
 AGCATGGACT GTATCCGCAT GCAGGACTCG GACCTGAGTG ACCCCATGTG GCCACAGTAC 360  
 ACGAACCTGG GGCTCCTGAA CAGCATGGAC CAGCAGATTC AGAACGGCTC CTCGTCCACC 420  
 AGTCCCCTATA ACACAGACCA CGCGCAGAAG AGCGTCACGG CGCCCTCGCC CTACGCACAG 480  
 CCCAGCTCCA CCTTCGATGC TCTCTCTCCA TCCCGGCCA TCCCTCCAA CACCGACTAC 540  
 55 CCAGGCCCGC ACAGTTTCGA CGTGTCTTTC CAGCAGTCGA GCACCGCCAA GTCGGCCACC 600  
 TGGACGTATT CCACTGAAC TGAAGAACTC TACTGCCAAA TTGCAAAGAC ATGCCCCATC 660  
 CAGATCAAGG TGATGACCCC ACCTCCTCAG GGAGCTGTTA TCCGCGCCAT GCCTGTCTAC 720  
 AAAAAAGCTG AGCACGTCAC GGAGGTGGTG AAGCGGTGCC CCAACCATGA GCTGAGCCGT 780  
 GAATTCACCG AGGGACAGAT TGCCCCCTCT AGTCATTGTA TTCGAGTAGA GGGGAACAGC 840  
 60 CATGCCCAGT ATGTAGAAGA TCCCACATCA GGAAGACAGA GTGTGCTGGT ACCTTATGAG 900  
 CCACCCCAAG TTGGCACTGA ATTACGACA GTCTTGTA TAATCATGTG TAACAGCAGT 960  
 TGTGTTGGAG GGATGAACCG CCGTCCAATT TTAATCATTG TTAATCTGGA AACCAGAGAT 1020  
 GGGCAAGTCC TGGGCCGACG CTGCTTTGAG GCCCGGATCT GTGCTTGCCC AGGAAGAGAC 1080  
 AGGAAGGCGG ATGAAGATAG CATCAGAAAG CAGCAAGTTT CGGACAGTAC AAAGAACGGT 1140  
 GATGGTACGA AGCGCCCGTT TCGTCAGAAC ACACATGGTA TCCAGATGAC ATCCATCAAG 1200  
 65 AAACGAAGAT CCCCAGATGA TGAAGCTTGA TACTTACCAG TGAGGGGCCG TGAGACTTAT 1260  
 GAAATGCTGT TGAAGATCAA AGAGTCCCTG GAACCTATGC AGTACCTTCC TCAGCACACA 1320  
 ATTGAAACGT ACAGGCAACA GCAACAGCAG CAGCACCAGC ACTTACTTCA GAAACATCTC 1380  
 CTTTCAGCCT GCTTCAGGAA TGAGCTTGTG GAGCCCCGGA GAGAACTCC AAAACAATCT 1440  
 70 GACGTCTTCT TTAGACATTG CAAGCCCCCA AACCAGTACG TGTACCCATA GAGCCCTATC 1500  
 TCTATATTTT AAGTGTGTGT GTTGTATTTT CATGTGTATA TGTGAGTGTG TGTGTGTGTA 1560  
 TGTGTGTGCG TGTGTATCTA GCCCTCATAA ACAGGACTTG AAGACACTTT GGCTCAGAGA 1620  
 CCCAAGTCTT CAAAGGCACA AAGCCACTAG TGAGGAATC TTTTGAAGGG ACTCAAACCT 1680  
 TTACAAGAAA GGATGTTTTT TGCAGATTTT GTATCCTTAG ACCGGCCATT GGTGGGTGAG 1740  
 75 GAACCACTGT GTTGTCTGTG GAGCTTTCTG TTGTTTCTCT GGAGGGAGGG GTCAGGTGGG 1800  
 GAAAGGGGCA TTAAGATGTT TATTGGAACC CTTTCTGTCT TTCTTCTGTT GTTTTTCTAA 1860  
 AATTCACAGG GAAGCTTTTG AGCAGGTCTC AAACCTAAGA TGTCTTTTAA AGAAAAGGAG 1920  
 AAAAAAGTTG TTATTGTCTG TGCATAAGTA AGTTGTAGGT GACTGAGAGA CTCAGTCAGA 1980  
 80 CCCTTTTAAT GCTGGTCATG TAATAATATT GCAAGTAGTA AGAAACGAAG GTGTCAAGTG 2040  
 TACTGCTGGG CAGCGAGGTG ATCATTACCA AAAGTAATCA ACTTTGTGGG TGGAGAGTTC 2100  
 TTTGTGAGAA CTTGCATTAT TTGTGTCTCT CCCTCATGTG TAGGTAGAAC ATTTCTTAAT 2160  
 GCTGTGTACC TGCTCTGCTC ACTGTATGTT GGCATCTGTT ATGCTAAAGT TTTTCTTGTA 2220  
 CATGAAACCC TGGAGACCTT ACTACAAAAA AACTGTTGTT TGGCCCCCAT AGCAGGTGAA 2280  
 CTCATTTTGT GCTTTTAATA GAAAGACAAA TCCACCCAG TAATATTGCC CTTACGTAGT 2340  
 85 TGTTTACCAT TAATCAAAGC TCAAAATAGA ATTTGAAGCC CTCTCACAAA ATCTGTGATT 2400  
 AATTTGCTTA ATTAGAGCTT CTATCCCTCA AGCCTACCTA CCATAAAACC AGCCATATTA 2460  
 CTGATACTGT TCAGTGCAAT TAGCCAGGAG ACTTACGTTT TGAGTAAGTG AGATCCAAGC 2520  
 AGACGTGTTA AAATCAGCAC TCCTGGACTG GAAATTAAAG ATTGAAAGGG TAGACTACTT 2580



TTCTTTTTTT TACTCAAAAG TTTAGAGAAT CTCTGTTTCT TTCCATTTTA AAAACATATT 2640  
 TTAAGATAAT AGCATAAAGA CTTTAAAAAT GTTCCTCCCC TCCATCTTCC CACACCCAGT 2700  
 CACCAGCACT GTATTTTCTG TCACCAAGAC AATGATTTCT TGTATTGAG GCTGTTGCTT 2760  
 TTGTGGATGT GTGATTTTAA TTTTCAATAA ACTTTTGCAT CTTGGTTTAA AAGAAA

Seq ID NO: 22 Protein sequence:  
 Protein Accession #: NP\_003713

1 11 21 31 41 51  
 | | | | | |  
 MSQSTQTNEF LSPEVFQHIW DFLEQPICSV QPIDLNFVDE PSEDGATNKI EISMDCIRMQ 60  
 DSDLSDPMWP QYTNLGLLNS MDQQIQNGSS STSPYNTDHA QNSVTAPSPY AQPSSSTFDAL 120  
 SPSPAIPSN T DYPGPHSFDV SFQSSSTAKS ATWTYSTELK KLYCQIAKTC PIQIKVMTTP 180  
 PQGAVIRAMP VYKKAHVTE VVKRCPNHEL SREFNEGQIA PPSHLIRVEG NSHAQYVEDP 240  
 ITGRQSVLVP YEPPQVGTET TTVLYNFMEN SSCVGGMNRR PILIIVTLET RDGQVLGRRR 300  
 FEARICACPG RDRKADEDSI RKQQVSDSTK NGDGTKRPFQ QNTHGIQMTS IKKRRSPDDE 360  
 LLYLPVRGRE TYEMLLKIKE SLELMQYLPQ HTIETIRQQQ QQQHQHLLQK HLLSACFRNE 420  
 LVEPRRETPK QSDVFFRHSK PPNRSVYP

Seq ID NO: 23 DNA sequence  
 Nucleic Acid Accession #: NM\_001944.1  
 Coding sequence: 84-3083

1 11 21 31 41 51  
 | | | | | |  
 TTTTCTTAGA CATTAAGTGC AGACGGCTGG CAGGATAGAA GCAGCGGCTC ACTTGGAATT 60  
 TTTTACCAGG GAAATCAGAG ACAATGATGG GGCTCTTCCC CAGAACTACA GGGGCTCTGG 120  
 CCATCTTCGT GGTGGTCATA TTGGTTCATG GAGAATTGCG AATAGAGACT AAAGGTCAAT 180  
 ATGATGAAGA AGAGATGACT ATGCAACAAG CTAAAAGAAG GCAAAAACGT GAATGGGTGA 240  
 AATTTGCCAA ACCCTGCAGA GAAGGAGAAG ATAAGTCAAA AAGAAACCCA ATTGCCAAGA 300  
 TTACTTCAGA TTACCAAGCA ACCCAGAAAA TCACCTACCG AATCTCTGGA GTGGGAATCG 360  
 ATCAGCCGCC TTTTGGGAATC TTTGTTGTTG ACAAAAACAC TGGAGATATT AACATAACAG 420  
 CTATAGTCGA CCGGAGGAGAA ACTCCAAGCT TCCTGATCAC ATGTCGGGCT CTAAATGCCC 480  
 AAGGACTAGA TGTAGAGAAA CCACTTATAC TAACGGTTAA AATTTTGGAT ATTAATGATA 540  
 ATCCTCCAGT ATTTTCACAA CAAATTTTCA TGGGTGAAAT TGAAGAAAAT AGTGCCTCAA 600  
 ACTCACTGGT GATGATACTA AATGCCACAG ATGCAGATGA ACCAAACCAC TTGAATTCTA 660  
 AAATTGCCTT CAAATTTGTC TCTCAGGAAC CAGCAGGCAC ACCCATGTTT CTCCTAAGCA 720  
 GAAACACTGG GGAAGTCCGT ACTTTGACCA ATTCTCTTGA CCGAGAGCAA GCTAGCAGCT 780  
 ATCGTCTGGT TGTGAGTGGT GCAGACAAAG ATGGAGAAGG ACTATCAACT CAATGTGAAT 840  
 GTAATATTAA AGTAAAGAT GTCAACGATA ACTTCCCAAT GTTAGAGAC TCTCAGTATT 900  
 CAGCACGTAT TGAAGAAAAT ATTTTAAGTT CTGAATTACT TCGATTTCAT GTAACAGATT 960  
 TGGATGAAGA GTACACAGAT AATTGGCTTG CAGTATATTT CTTTACCTCT GGAATGAAG 1020  
 GAAATTGGTT TGAATACAAA ACTGATCCTA GAACTAATGA AGGCATCCTG AAAGTGGTGA 1080  
 AGGCTCTAGA TTATGAACAA CTACAAAGCG TGAACTTAG TATTGCTGTC AAAAACAAG 1140  
 CTGAATTTCA CCAATCAGTT ATCTCTCGAT ACCGAGTTCA GTCAACCCCA GTCACAATTC 1200  
 AGTAATATAA TGTAAAGAGAA GGAATTGCAT TCCGTCTCTG TCCCAAGACA TTTACTGTGC 1260  
 AAAAAGGCAT AAGTAGCAAA AAATTGGTGG ATTATATCCT GGGAACATAT CAAGCCATCG 1320  
 ATGAGGACAC TAACAAAGCT GCCTCAAATG TCAAATATGT CATGGGACGT AACGATGGTG 1380  
 GATACCTAAT GATTGATTCA AAAACTGCTG AAATCAAATT TGTCAAAAAT ATGAACCGAG 1440  
 ATTCTACTTT CATAGTTAAC AAAACAATCA CAGCTGAGGT TCTGGCCATA GATGAATACA 1500  
 CGGGTAAAGC TTCTACAGGC ACGGTATATG TTAGATTACC CGATTTCAT GACAATTGTC 1560  
 CAACAGCTGT CCTCGAAAAA GATGCAGTTT GCGATTCTTC ACCTTCCGTG GTTGTCTCCG 1620  
 CTAGAACACT GAATAATAGA TACACTGGCC CCTATACATT TGCACTGGAA GATCAACCTG 1680  
 TAAAGTTGCC TGCCGTATGG AGTATCACA CCCTCAATGC TACCTCGGCC CTCCTCAGAG 1740  
 CCCAGGAACA GATACCTCCT GGAGTATACC ACATCTCCCT GGTACTTACA GACAGTCAGA 1800  
 ACAATCGGTG TGAGATGCCA CGCAGCTTGA CACTGGAAGT CTGTCACTGT GACAACAGGG 1860  
 GCATCTGTGG AACTTCTTAC CCAACCAAGA GCCCTGGGAC CAGGTATGGC AGGCCGCACT 1920  
 CAGGAGAGCT GGGGCTGCC GCCATCGCC TGTCTCTCCT TGGTCTCCTG CTGCTGCTGT 1980  
 TGGCCCCCCT TCTGCTGTTG ACCTGTGACT GTGGGGCAGG TTCTACTGGG GGAGTGACAG 2040  
 GTGGTTTTAT CCCAGTTCTT GATGGCTCAG AAGGAACAAT TCATCAGTGG GGAATTGAAG 2100  
 GAGCCCATCC TGAAGACAAG GAAATCACA ATATTTGTGT GCCTCCTGTA ACAGCCAATG 2160  
 GAGCCGATTT CATGGAAGT TCTGAAGTTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220  
 TGGAAGGCAC TTCAGGAATG GAAATGACCA CTAAGCTTGG AGCAGCCACT GAATCTGGAG 2280  
 GTGCTGCAGG CTTTGCAACA GGGACAGTGT CAGGAGCTGC TTCAGGATTC GGAGCAGCCA 2340  
 CTGGAGTTGG CATCTGTTCC TCAGGGCAGT CTGGAACCAT GAGAACAAGG CATTCCACTG 2400  
 GAGGAACCAA TAAGGACTAC GCTGATGGGG CGATAAGCAT GAATTTCTG GACTCCTACT 2460  
 TTTCTCAGAA AGCATTTGCC TGTGCGGAGG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520  
 TGTGATCTA TGATAATGAA GGCAGAGATG CCACTGGTTC TCCTGTGGGC TCCGTGGGTT 2580  
 GTTGCACTTT TATTGCTGAT GACCTGGATG ACAGCTTCTT GGACTCACTT GGACCCAAAT 2640  
 TTAATAAAGT TGCAGAGATA AGCCTTGGTG TTGATGGTGA AGGCAAGAA GTTCAGCCAC 2700  
 CCTCTAAAGA CAGCGGTTAT GGGATTGAAT CCTGTGGCCA TCCCATAGAA GTCCAGCAGA 2760  
 CAGGATTTGT TAAGTGCCAG ACTTTGTGAG GAAGTCAAGG AGCTTCTGCT TTGTCCGCCT 2820  
 CTGGGTCTGT CCAGCCAGCT GTTTCCATCC CTGACCCTCT GCAGCATGGT AACTATTTAG 2880  
 TAACGGAGAC TTACTCGGCT TCTGGTTCCC TCGTGCAACC TTCCACTGCA GGCTTTGATC 2940  
 CACTTCTCAC ACAAATGTG ATAGTGACAG AAAGGGTGAT CTGTCCCATT TCCAGTGTTC 3000  
 CTGGCAACCT AGCTGGCCCA ACGCAGCTAC GAGGGTCACA TACTATGCTC TGTACAGAGG 3060  
 ATCCTTGCTC CCGTCTAATA TGACCAAGAT GAGCTGGAAT ACCCACTGTA CCAAACTGCG 3120  
 ATCTTTGGAC TAAAGTATTC AAAATAGCAT AGCAAAGCTC ACTGTATTGG GCTAATAATT 3180  
 TGGCACTTAT TAGCTTCTCT CATAAACTGA TCACGATTAT AAATTAAATG TTTGGGTTCA 3240  
 TACCCCAAAA GCAATATGTT GTCACTCCTA ATTCTCAAGT ACTATTCAAA TTGTAGTAAA 3300  
 TCTTAAAGTT TTTCAAAACC CTAAATCAT ATTCGC

Seq ID NO: 24 Protein sequence:  
 Protein Accession #: NP\_001935.1

1 11 21 31 41 51



	MMGLFPRTTG	ALAIFVVVIL	VHGELEIETK	GQYDEEEMTM	QQAARRQKRE	WVKFAKPCRE	60
	GEDNSKRNPI	AKITSDYQAT	QKITYRISGV	GIDQPPFGIF	VVDKNTGDIN	ITAIVDREET	120
5	PSFLITCRAL	NAQGLDVEKP	LILTVKILDI	NDNPPVFSQQ	IFMGEIEENS	ASNSLVMILN	180
	ATDADEPNHL	NSKIAFKIVS	QEPAGTPMFL	LSRNTGEVRT	LTNSLDREQA	SSYRLVVSQA	240
	DKDGEGLSTQ	CECNIKVKDV	NDNFPMFRDS	QYSARIEENI	LSSELLRFQV	TDLDEEYTDN	300
	WLAVYFFTSQ	NEGNWFIEQT	DPRTNEGILK	VVKALDYEQL	QSVKLSIAVK	NKAEFHQSVI	360
	SRYRVQSTPV	TIQVINVREG	IAFRPASKTF	TVQKGISSKK	LVDYILGTQY	AIDEDTNKAA	420
10	SNVKYVMGRN	DGGYLMIDSK	TAEIKFVKNM	NRDSTFIVNK	TITAEVLAIQ	EYTGKTSTGT	480
	VYVRVPDFND	NCPTAVLEKD	AVCSSSPSVV	VSARTLNNRY	TGPYTFALQD	QPVKLPVWS	540
	ITTLNATSAL	LRAQEQIPPG	VYHISLVLTQ	SQNNRCCEMPR	SLTLEVCQCD	NRGICGTSYP	600
	TTSPGTRYGR	PHSGRLGPAA	IGLLLLGLLL	LLLAPLLLLT	CDGAGSTGG	VTGGFIPVPD	660
	GSEGTIHQWG	IEGAHPEDKE	ITNICVPPVT	ANGADFMESQ	EVCTNTYARG	TAVEGTSQME	720
15	MTTKLGAATE	SGGAAGFATG	TVSGAASGFG	AATGVGICSS	GQSGTMRTRH	STGGTNKDYA	780
	DGAISMNFLD	SYFSQKAFAC	AEEDDGQEAN	DCLLIYDNEG	ADATGSPVGS	VGCCSFIADD	840
	LDDSFLLSLG	PKFKKLAEIS	LGVDGEGKEV	QPPSKDSGYG	IESCGHPIEV	QQTGFVKCQT	900
	LSGSQASAL	SASGSVQPAV	SIPDPLQHGN	YLVTTETYSAS	GSLVQPSTAG	FDPLLTQNVQ	960
	VTERTVICPIS	SVPGNLAGPT	QLRGSHTMLC	TEDPCSRLLI			

Seq ID NO: 25 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 56-1642

25	1	11	21	31	41	51	
	AGTATCCCAG	GAGGAGCAAG	TGGCACGTCT	TCGGACCTAG	GCTGCCCTTG	CCGTCTATGTC	60
	GCAAGGGATC	CTTTCTCCGC	CAGCGGGCTT	GCTGTCCGAT	GACGATGTCT	TAGTTTCTCC	120
	CATGTTTGAG	TCCACAGCTG	CAGATTGGGG	GTCTGTGGTA	CGCAAGAACC	TGCTATCAGA	180
	CTGCTCTGTC	GTCTCTACCT	CCCTAGAGGA	CAAGCAGCAG	GTTCCATCTG	AGGACAGTAT	240
30	GGAGAAGGTG	AAAGTATACT	TGAGGGTTAG	GCCCTTGTTA	CCTTCAGAGT	TGGAACGACA	300
	GGAAGATCAG	GGTTGTGTCC	GTATTGAGAA	TGTGGAGACC	CTTGTCTTAC	AAGCACCCAA	360
	GGACTCTTTT	GCCCTGAAGA	GCAATGAACG	GGGAATTGGC	CAAGCCACAC	ACAGGTTTAC	420
	CTTTTCCCAG	ATCTTTGGGC	CAGAAGTGGG	ACAGGCATCC	TTCTTCAACC	TAACTGTGAA	480
35	GGAGATGGTA	AAGGATGTAC	TCAAAGGGCA	GAAGTGGCTC	ATCTATACAT	ATGGAGTCAC	540
	TAACTCAGGG	AAAACCCACA	CGATTCAAGG	TACCATCAAG	GATGGAGGGA	TTCTCCCCCG	600
	GTCCCTGGCG	CTGATCTTCA	ATAGCCTCCA	AGGCCAACTT	CATCCAACAC	CTGATCTGAA	660
	GCCCTTGCTC	TCCAATGAGG	TAATCTGGCT	AGACAGCAAG	CAGATCCGAC	AGGAGGAAAT	720
	GAAGAAGCTG	TCCCTGCTAA	ATGGAGGCCT	CCAAGAGGAG	GAGCTGTCCA	CTTCCTTGAA	780
40	GAGGAGTGTC	TACATCGAAA	GTCCGATAGG	TACCAGCACC	AGCTTCGACA	GTGGCATTGC	840
	TGGGCTCTCT	TCTATCAGTC	AGTGATACCAG	CAGTAGCCAG	CTGGATGAAA	CAAGTCATCG	900
	ATGGGCACAG	CCAGACACTG	CCCCACTACC	TGTCCCGGCA	AACATTGCTC	TCTCCATCTG	960
	GATCTCATTC	TTTGAGATCT	ACAACGAACT	GCTTTATGAC	CTATTAGAAC	CGCCTAGCCA	1020
	ACAGCGCAAG	AGGCAGACTT	TGCGGCTATG	CGAGGATCAA	AATGGCAATC	CCTATGTGAA	1080
45	AGATCTCAAC	TGGATTCAAT	TGCAAGATGC	TGAGGAGGCC	TGGAAGCTCC	TAAAAGTGGG	1140
	TCGTAAGAAC	CAGAGCTTTG	CCAGCACCCA	CCTCAACCAG	AATCCAGCC	GCAGTCACAG	1200
	CATCTTCTCA	ATCAGGATCC	TACACCTTCA	GGGGGAAGGA	GATATAGTCC	CCAAGATCAG	1260
	CGAGCTGTCA	CTCTGTGATC	TGGCTGGCTC	AGAGCGCTGC	AAAGATCAGA	AGAGTGGTGA	1320
	ACGGTTGAAG	GAAGCAGGAA	ACATTAACAC	CTCTCTACAC	ACCCTGGGCC	GCTGTATTGC	1380
50	TGCCCTTCGT	CAAAACCAGC	AGAACCGGTC	AAAGCAGAAC	CTGGTTCCCT	TCCGTGACAG	1440
	CAAGTTGACT	CGAGTGTTCC	AAGGTTTCTT	CACAGGCCGA	GGCCGTTCCT	GCATGATTGT	1500
	CAATGTGAAT	CCCTGTGCAT	CTACCTATGA	TGAAACTCTT	CATGTGGCCA	AGTTCTCAGC	1560
	CATTGTAGC	CAGGTGACTT	GTGCATGCC	CACCTATGCA	ACTGGGATTC	CCATCCCTGC	1620
	ACTCGTTCAT	CAAGGAACAT	AGTCTTCAGG	TATCCCCCAG	CTTAGAGAAA	GGGGCTAAGG	1680
55	CAGACACAGG	CCTTGATGAT	GATATTGAAA	ATGAAGCTGA	CATCTCCATG	TATGGCAAAG	1740
	AGGAGCTCCT	ACAAGTTGTG	GAAGCCATGA	AGACACTGCT	TTTGAAGGAA	CGACAGGAAA	1800
	AGCTACAGCT	GGAGATGCAT	CTCCGAGATG	AAATTTGCAA	TGAGATGGTA	GAACAGATGC	1860
	AACAGCGGGA	ACAGTGGTGC	AGTGAACATT	TGGACACCCA	AAAGGAACTA	TTGGAGGAAA	1920
	TGTATGAAGA	AAAATAAAT	ATCCTCAAGG	AGTCACTGAC	AAGTTTTTAC	CAAGAAGAGA	1980
60	TTCAGGAGCG	GGATGAAAAG	ATTGAAGAGC	TAGAAGCTCT	CTTGCAGGAA	GCCAGACAAC	2040
	AGTCAGTGGC	CCATCAGCAA	TCAGGCTCTG	AATTGGCCCT	ACGGCGGTCA	CAAAGGTTGG	2100
	CAGCTTCTGC	CTCCACCCAG	CAGCTTCAGG	AGGTTAAAGC	TAAATTACAG	CAGTGCAAAG	2160
	CAGAGCTAAA	CTCTACCACT	GAAGAGTTGC	ATAAGTATCA	GAAAATGTTA	GAACCACCAC	2220
	CCTCAGCCAA	GCCCTTCACC	ATTGATGTGG	ACAAGAAGTT	AGAAGAGGGC	CAGAAGAATA	2280
65	TAAGGCTGTT	GCGGACAGAG	CTTCAGAAAC	TTGGTGAGTC	TCTCCAATCA	GCAGAGAGAG	2340
	CTTGTTGCCA	CAGCACTGGG	GCAGGAAAAC	TTCTGCAAGC	CTTGACCACT	TGTGATGACA	2400
	TCTTAATCAA	ACAGGACCAG	ACTCTGGCTG	AAGTCAGAAA	CAACATGGTG	CTAGTGAAAC	2460
	TGGACCTTCG	GAAGAAGGCA	GCATGTATTG	CTGAGCAGTA	TCATACTGTG	TTGAAACTCC	2520
	AAGGCCAGGT	TTCTGCCAAA	AAGCGCCTTG	GTACCAACCA	GGAAAATCAG	CAACCAAACC	2580
70	AACAACCACC	AGGGAAGAAA	CCATTCCCTC	GAAATTTACT	TCCCGAACA	CCAACCTGCC	2640
	AAAGCTCAAC	AGACTGCAGC	CCTTATGCCC	GGATCCTACG	CTCACGGCGT	TCCCCTTTAC	2700
	TCAAATCTGG	GCCTTTTGGC	AAAAAGTACT	AAGGCTGTGG	GGAAAGAGAA	GAGCAGTCAT	2760
	GGCCCTGAGG	TGGGTCAGCT	ACTCTCCTGA	AGAAATAGGT	CTCTTTTATG	CTTTACCATA	2820
	TATCAGGAAT	TATATCCAGG	ATGCAATACT	CAGACACTAG	CTTTTCTCTC	ACTTTTGTAT	2880
75	TATAACCACC	TATGTAATCT	CATGTTGTTG	TTTTTTTTTA	TTTACTTATA	TGATTTCTAT	2940
	GCACACAAAA	ACAGTTATAT	TAAAGATATT	ATTGTTTACA	TTTTTTATTG	AATTCCAAAT	3000
	GTAGCAAAAT	CATTAAACA	AATTATAAAA	GGGACAGAAA	AA		

Seq ID NO: 26 Protein sequence:  
Protein Accession #: Eos sequence

80	1	11	21	31	41	51	
	MSQGILSPPA	GLLSDDDVVV	SPMFESTAAD	LGSVVRKNLL	SDCSVVSTSL	EDKQVPSQD	60
	SMEKVKVYLR	VRPLLPSELE	RQEDQGCVRQ	ENVETLVLLQ	PKDSFALKSN	ERGIGQATHR	120
85	FTFSQIFGPE	VGQASFFNLT	VKEMVKDVLK	GQNWLIYTYG	VTNSGKTHTI	QGTIKDGGIL	180
	PRSLALIFNS	LQGLHPTPD	LKPLLSNEVI	WLDSKQIRQE	EMKKLSLLNG	GLQEEELSTS	240
	LKRSVYIESR	IGTSTSFDSG	IAGLSSISQC	TSSSQLDETS	HRWAQPDATP	LPVPANIRFS	300



IWISFFEIYN ELLYDLLEPP SQQRKRQTLR LCEDQNGNPY VKDLNWIHVQ DAEEAWKLLK 360  
 VGRKNQSFAS THLNQNSSRS HSIFSIRILH LQEGEDIVPK ISELSLCDLA GSERCKDQKS 420  
 GERLKEAGNI NTSLHTLGRC IAALRQNQQN RSKQNLVPFR DSKLTRVFGG FFTGRGRSCM 480  
 IVNVNPCAST YDETLHVAKF SAIASQVTCA CPTYATGIPI PALVHQGT

Seq ID NO: 27 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 13-1424

1	11	21	31	41	51	
TAGAAGTTTA	CAATGAAGTT	TCTTCTAATA	CTGCTCCTGC	AGGCCACTGC	TTCTGGAGCT	60
CTTCCCCTGA	ACAGCTCTAC	AAGCCTGGAA	AAAAATAATG	TGCTATTTGG	TGAAAGATAC	120
TTAGAAAAAT	TTTATGGCCT	TGAGATAAAC	AAACTTCCAG	TGACAAAAAT	GAAATATAGT	180
GGAAACTTAA	TGAAGGAAAA	AATCCAAGAA	ATGCAGCACT	TCTTGGGTCT	GAAAGTGACC	240
GGGCAACTGG	ACACATCTAC	CCTGGAGATG	ATGCACGCAC	CTCGATGTGG	AGTCCCCGAT	300
GTCCATCATT	TCAGGGAAAT	GCCAGGGGGG	CCCGTATGGA	GGAAACATTA	TATCACCTAC	360
AGAATCAATA	ATTACACACC	TGACATGAAC	CGTGAGGATG	TTGACTACGC	AATCCGGAAA	420
GCTTTCCAAG	TATGGAGTAA	TGTTACCCCC	TTGAAATTCA	GCAAGATTAA	CACAGGCATG	480
GCTGACATTT	TGGTGGTTTT	TGCCCCGTGA	GCTCATGGAG	ACTTCCATGC	TTTTGATGGC	540
AAAGGTGGAA	TCCTAGCCCA	TGCTTTTGGA	CCTGGATCTG	GCAATTGGAG	GGATGCACAT	600
TTCGATGAGG	ACGAATTCTG	GACTACACAT	TCAGGAGGCA	CAAACTTGTT	CCTCACTGCT	660
GTTCACGAGA	TTGGCCATTC	CTTAGGTCTT	GGCCATTCTA	GTGATCCAAA	GGCCGTAATG	720
TTCCCCACCT	ACAAATATGT	TGACATCAAC	ACATTTTCGCC	TCTCTGCTGA	TGACATACGT	780
GGCATTCACT	CCCTGTATGG	AGACCCAAAA	GAGAACCAC	GCTTGCCAAA	TCCTGACAAT	840
TCAGAACCAG	CTCTCTGTGA	CCCCAATTTG	AGTTTTGATG	CTGTCACTAC	CGTGGGAAAT	900
AAGATCTTTT	TCTTCAAAGA	CAGGTTCTTC	TGGCTGAAGG	TTTCTGAGAG	ACCAAAGACC	960
AGTGTTAATT	TAATTTCTTC	CTTATGGCCA	ACCTTGCCAT	CTGGCATTGA	AGCTGCTTAT	1020
GAAATTGAAG	CCAGAAATCA	AGTTTTTCTT	TTTAAAGATG	ACAAATACTG	GTTAATTAGC	1080
AATTTAAGAC	CAGAGCCAAA	TTATCCCAAG	AGCATACTAT	CTTTTGGTTT	TCCTAACTTT	1140
GTGAAAAAAA	TTGATGCAGC	TGTTTTTAAC	CCACGTTTTT	ATAGGACCTA	CTTCTTTGTA	1200
GATAACCAGT	ATTGGAGGTA	TGATGAAAGG	AGACAGATGA	TGGACCCTGG	TTATCCCAAA	1260
CTGATTACCA	AGAACTTCCA	AGGAATCGGG	CCTAAAATTG	ATGCAGTCTT	CTACTCTAAA	1320
AACAAATACT	ACTATTCTTT	CCAAGGATCT	AACCAATTG	AATATGACTT	CCTACTCCAA	1380
CGTATCACCA	AAACACTGAA	AAGCAATAGC	TGGTTTGGTT	GTTGAAAATG	GTGTAATTAA	1440
TGGTTTTTGT	TAGTTCACCT	CAGCTTAATA	AGTATTTAT	GCATATTTGC	TATGTCCTCA	1500
GTGTACCACT	ACTTAGAGAT	ATGTATCATA	AAAATAAAAT	CTGTAAACCA	TAGGTAATGA	1560
TTATATAAAA	TACATAATAT	TTTTCAATTT	TGAAAACCT	AATTGTCCAT	TCTTGCTTGA	1620
CTCTACTATT	AAGTTTGAAA	ATAGTTACCT	TCAAAGCAAG	ATAATTCTAT	TTGAAGCATG	1680
CTCTGTAAGT	TGCTTCCTAA	CATCCTTGGA	CTGAGAAATT	ATACTTACTT	CTGGCATAAC	1740
TAAAATTAAG	TATATATATT	TTGGCTCAAA	TAAAATTG			

Seq ID NO: 28 Protein sequence:  
 Protein Accession #: Eos sequence

1	11	21	31	41	51	
MKFLILLILLQ	ATASGALPLN	SSTSLEKNNV	LFGERYLEKF	YGLEINKLPV	TKMKYSGNLM	60
KEKIQEMQHF	LGLKVTGQLD	TSTLEMMHAP	RCGVDPDVHF	REMPGGPVWR	KHYITYRINN	120
YTPDMNREDV	DYAIRKAFQV	WSNVTPPKFS	KINTGMADIL	VVFARGAHGD	FHAFDGGKGI	180
LAHAFPGPSG	IGGDAHFDED	EFWTTTHSGG	NLFLTAVHEI	GHSLSGLGHS	DPKAVMFPTY	240
KYVDINTFRL	SADDIRGIQS	LYGDPKENQR	LPNPDNSEPA	LCDPNLSFDA	VTTVGNKIFF	300
FKDRFFWLKV	SERPKTSVNL	ISSLWPTLPS	GIEAAYEIEA	RNQVFLFKDD	KYWLISNLRP	360
EPNYPKSIHS	FGFPNFVKKI	DAAVFNPRFY	RTYFFVDNQY	WRYDERRQMM	DPGYPKLITK	420
NFQGIGPKID	AVFYSKNKYY	YFFQGSNQFE	YDFLLQRITK	TLKSNSWFGC		

Seq ID NO: 29 DNA sequence  
 Nucleic Acid Accession #: NM\_006115.1  
 Coding sequence: 236..1765

1	11	21	31	41	51	
GCTTCAGGGT	ACAGCTCCCC	CGCAGCCAGA	AGCCGGGCCT	GCAGCCCCTC	AGCACCGCTC	60
CGGGACACCC	CACCCGCTTC	CCAGGCGTGA	CCTGTCAACA	GCAACTTCGC	GGTGTGGTGA	120
ACTCTCTGAG	GAAAAACCAT	TTTGATTATT	ACTGTCAGAC	GTGCGTGGCA	ACAAGTGACT	180
GAGACCTAGA	AATCCAAGCG	TTGGAGGTCC	TGAGGCCAGC	CTAAGTCGCT	TCAAAATGGA	240
ACGAAGGCGT	TTGTGGGGTT	CCATTCAAG	CCGATACATC	AGCATGAGTG	TGTGGACAAG	300
CCCACGGAGA	CTTGTGGAGC	TGGCAGGGCA	GAGCCTGCTG	AAGGATGAGG	CCCTGGCCAT	360
TGCCGCCCTG	GAGTTGCTGC	CCAGGGAGCT	CTTCCCGCCA	CTCTTCATGG	CAGCCTTTGA	420
CGGGAGACAC	AGCCAGACCC	TGAAGGCAAT	GGTGCAGGCC	TGGCCCTTCA	CCTGCCTCCC	480
TCTGGGAGTG	CTGATGAAGG	GACAACATCT	TCACCTGGAG	ACCTTCAAAG	CTGTGCTTGA	540
TGGACTTGAT	GTGCTCCTTG	CCCAGGAGGT	TCGCCCCAGG	AGGTGGAAAC	TTCAAGTGCT	600
GGATTTACGG	AAGAACTCTC	ATCAGGACTT	CTGGACTGTA	TGGTCTGGAA	ACAGGGCCAG	660
TCTGTACTCA	TTTCCAGAGC	CAGAAGCAGC	TCAGCCCATG	ACAAAGAAGC	GAAAAGTAGA	720
TGGTTTGAGC	ACAGAGGCAG	AGCAGCCCTT	CATTCCAGTA	GAGGTGCTCG	TAGACCTGTT	780
CCTCAAGGAA	GGTGCTGTG	ATGAATTGTT	CTCCTACCTC	ATTGAGAAAG	TGAAGCGAAA	840
GAAAAATGTA	CTACGCCTGT	GCTGTAAGAA	GCTGAAGATT	TTTGCAATGC	CCATGCAGGA	900
TATCAAGATG	ATCCTGAAAA	TGGTGCAGCT	GGACTCTATT	GAAGATTGG	AAGTGACTTG	960
TACCTGGAAG	CTACCCACCT	TGGCGAAATT	TTCTCCTTAC	CTGGGCCAGA	TGATTAATCT	1020
GCGTAGACTC	CTCCTCTCCC	ACATCCATGC	ATCTTCCTAC	ATTTCCCCGG	AGAAGGAAGA	1080
GCAGTATATC	GCCCAGTTCA	CCTCTCAGTT	CCTCAGTCTG	CAGTGCCTGC	AGGCTCTCTA	1140
TGTGGACTCT	TTATTTTTC	TTAGAGGCCG	CCTGGATCAG	TTGCTCAGGC	ACGTGATGAA	1200
CCCCTTGGA	ACCTCTCAA	TAACCTAAGT	CCGGCTTTTC	GAAGGGGATG	TGATGCATCT	1260
GTCCCAGAGT	CCCAGCGTCA	GTCAGTCAAG	TGTCCTGAGT	CTAAGTGGGG	TCATGCTGAC	1320
CGATGTAAGT	CCGAGCCCC	TCCAAGCTCT	GCTGGAGAGA	GCCTCTGCCA	CCCTCCAGGA	1380
CCTGGTCTTT	GATGAGTGTG	GGATCACGGA	TGATCAGCTC	CTTGCCCTCC	TGCCTTCCCT	1440
GAGCCACTGC	TCCCAGCTTA	CAACCTTAAG	CTTCTACGGG	AATTCCATCT	CCATATCTGC	1500



5 CTTGCAGAGT CTCCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCACG TGCTGTATCC 1560  
 TGTCCCCCTG GAGAGTTATG AGGACATCCA TGGTACCCCTC CACCTGGAGA GGCTTGCCTA 1620  
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCAGCA TGGTCTGGCT 1680  
 TAGTGCCAAC CCCTGTCCCTC ACTGTGGGGA CAGAACCTTC TATGACCCGG AGCCCATCCT 1740  
 GTGCCCCTGT TTCATGCCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800  
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTTGAAGCAA CAAAGCAGCC ACAGTTTCAG 1860  
 ACAAATGTTT AGTGTGAGTG AGGAAAACAT GTTCAGTGAG GAAAAACAT TCAGACAAAT 1920  
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTTGACTTG AGGAGTTAAT 1980  
 10 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040  
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAAGAAAC 2100  
 TGTGAAAAT AAAGAGAAGC AATGTGAAGC AAAAAAAAAA AAAAAAAA

Seq ID NO: 30 Protein sequence:  
 Protein Accession #: NP\_006106.1

15 1 11 21 31 41 51  
 | | | | | |  
 GCTTCAGGGT ACAGCTCCCC CGCAGCCAGA AGCCGGGCCT GCAGCGCCTC AGCACCGCTC 60  
 CGGGACACCC CACCCGCTTC CCAGGCGTGA CCTGTCAACA GCAACTTCGC GGTGTGGTGA 120  
 20 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGAAT 180  
 GAGACCTAGA AATCCAAGCG TTGGAGGTCC TGAGGCCAGC CTAAGTCGCT TCAAAATGGA 240  
 ACGAAGGCGT TTGTGGGGTT CCATTGAGAG CCGATACATC AGCATGAGTG TGTGGACAAG 300  
 CCCACGGAGA CTTGTGGAGC TGGCAGGGCA GAGCCTGCTG AAGGATGAGG CCCTGGCCAT 360  
 25 TGCCGCCCTG GAGTTGCTGC CCAGGGAGCT CTTCCCGCCA CTCTTCATGG CAGCCTTTGA 420  
 CGGGAGACAC AGCCAGACCC TGAAGGCAAT GGTGCAGGCC TGGCCCTTCA CCTGCCTCCC 480  
 TCTGGGAGTG CTGATGAAGG GACAACATCT TCACCTGGAG ACCTTCAAAG CTGTGCTTGA 540  
 TGGACTTGAT GTGCTCCTTG CCCAGGAGGT TCGCCCCAGG AGGTGGAAC TTCAAGTGCT 600  
 GGATTTACGG AAGAACTCTC ATCAGGACTT CTGGACTGTA TGGTCTGGAA ACAGGGCCAG 660  
 TCTGTACTCA TTTCCAGAGC CAGAAGCAGC TCAGCCCATG ACAAAGAAGC GAAAAGTAGA 720  
 30 TGGTTTGAGC ACAGAGGCAG AGCAGCCCTT CATTCCAGTA GAGGTGCTCG TAGACCTGTT 780  
 CCTCAAGGAA GGTGCCTGTG ATGAATTGTT CTCCTACCTC ATTGAGAAAG TGAAGCGAAA 840  
 GAAAAATGTA CTACGCCTGT GCTGTAAGAA GCTGAAGATT TTTGCAATGC CCATGCAGGA 900  
 TATCAAGATG ATCCTGAAAA TGGTGCAGCT GGACTCTATT GAAGATTGG AAGTGACTTG 960  
 35 TACCTGGAAG CTACCCACCT TGGCGAAATT TTCTCCTTAC CTGGGCCAGA TGATTAATCT 1020  
 GCGTAGACTC CTCCTCTCCC ACATCCATGC ATCTTCCTAC ATTTCCCCGG AGAAGGAAGA 1080  
 GCAGTATATC GCCCAGTTCA CCTCTCAGTT CCTCAGTCTG CAGTGCCTGC AGGCTCTCTA 1140  
 TGTGGACTCT TTATTTTTC TTAGAGGCCG CCTGGATCAG TTGCTCAGGC ACGTGATGAA 1200  
 CCCCTTGGA ACCCTCTCAA TAACTAACTG CCGGCTTTTC GAAGGGGATG TGATGCATCT 1260  
 40 GTCCAGAGT CCCAGCGTCA GTCAGCTAAG TGTCTGAGT CTAAGTGGGG TCATGCTGAC 1320  
 CGATGTAAGT CCCGAGCCCC TCCAAGCTCT GCTGGAGAGA GCCTCTGCCA CCCTCCAGGA 1380  
 CCTGGTCTTT GATGAGTGTG GGATCACGGA TGATCAGCTC CTTGCCCTCC TGCCTTCCCT 1440  
 GAGCCACTGC TCCCAGCTTA CAACCTTAAG CTTCTACGGG AATTCCATCT CCATATCTGC 1500  
 CTTGCAGAGT CTCCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCACG TGCTGTATCC 1560  
 45 TGTCCCCCTG GAGAGTTATG AGGACATCCA TGGTACCCCTC CACCTGGAGA GGCTTGCCTA 1620  
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCAGCA TGGTCTGGCT 1680  
 TAGTGCCAAC CCCTGTCCCTC ACTGTGGGGA CAGAACCTTC TATGACCCGG AGCCCATCCT 1740  
 GTGCCCCTGT TTCATGCCTA ACTAGCTGGG TGACATATC AAATGCTTCA TTCTGCATAC 1800  
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTTGAAGCAA CAAAGCAGCC ACAGTTTCAG 1860  
 50 ACAAATGTTT AGTGTGAGTG AGGAAAACAT GTTCAGTGAG GAAAAACAT TCAGACAAAT 1920  
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTTGACTTG AGGAGTTAAT 1980  
 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040  
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAAGAAAC 2100  
 TGTGAAAAT AAAGAGAAGC AATGTGAAGC AAAAAAAAAA AAAAAAAA

Seq ID NO: 31 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 64-2754

60 1 11 21 31 41 51  
 | | | | | |  
 GGCAGGTCTC GCTCTCGGCA CCCTCCCGGC GCCCGCGTTC TCCTGGCCCT GCCCGGCATC 60  
 CCGATGGCCG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120  
 65 CTGACCCTCG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180  
 CCTTCTAAAC TAGAGGCAGA CAAAATAATT GGCAGGTTA ATTTGGAAGA GTGCTTCAGG 240  
 TCTGCAGACC TCATCCGGTC AAGTGATCCT GATTTTCAGAG TTCTAAATGA TGGGTCAGTG 300  
 TACACAGCCA GGGCTGTTGC GCTGTCTGAT AAGAAAAGAT CATTATACCAT ATGGCTTTCT 360  
 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420  
 70 TCGAAGACAA GACACACTAG AGAAACTGTT CTGAGGCGTG CCAAGAGGAG ATGGGCACCT 480  
 ATTCCCTTGT CTATGCAAGA GAATTCCTTG GGCCCTTTCC CATTGTTTCT TCAACAAGTT 540  
 GAATCTGATG CAGCACAGAA CTATACTGTC TTCTACTCAA TAAGTGGACG TGGAGTTGAT 600  
 AAAGAACCCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG 660  
 CCTGTGGATC GTGAAGAATA TGATGTTTTT GATTTGATTG CTTATGCGTC AACTGCAGAT 720  
 GGATATTCAG CAGATCTGCC CCTCCCCTA CCCATCAGGG TAGAGGATGA AAATGACAAC 780  
 75 CACCCTGTTT TCACAGAAGC AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840  
 ACTACAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGCGC 900  
 CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACCTG GGCTCTTTTC TGTGCATCCC 960  
 AGCACAGGCG TAATCACCAC AGTCTCTCAT TATTTGGACA GAGAGGTTGT AGACAAGTAC 1020  
 TCATTGATAA TGAAAGTACA AGACATGGAT GGCCAGTTTT TTGGATTGAT AGGCACATCA 1080  
 80 ACTTGATCA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACCTTCAG ACAAATGCT 1140  
 TATGAAGCAT TTGTAGAGGA AAATGCATTC AATGTGGAAA TCTTACGAAT ACCTATAGAA 1200  
 GATAAGGATT TAATTAACAC TGCCAATTGG AGAGTCAATT TTACATTTT AAAGGGAAT 1260  
 GAAAATGGAC ATTTCAAAAT CAGCACAGAC AAAGAACTA ATGAAGGTGT TCTTTCTGTT 1320  
 85 GTAAAGCCAC TGAATTATGA AGAAAACCGT CAAGTGAACC TGGAAATTGG AGTAAACAAT 1380  
 GAAGCCCAT TTGCTAGAGA TATCCAGAG GTGACAGCCT TGAACAGAGC CTTGGTTACA 1440  
 GTTCATGTGA GGGATCTGGA TGAGGGGCCT GAATGCACTC CTGCAGCCCA ATATGTGCGG 1500  
 ATTAAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA 1560



	AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCAAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
5	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAAATGATA	ATCCACCAGA	AATACTTCAA	1800
	GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAGTTTGC	CCAATACTTC	TCCAGAAAATC	1920
	AGTAGACTGT	GGAGCCTCAC	CAAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGAATA	TACCATTCCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
10	GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACCTCA	GTGTCGTGCG	2100
	ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCTTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTTTGG	TGCAACTAAA	2220
	GGGAAACGTT	TTCCTGAAGA	TTTAGCACAG	CAAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
	CCTGGAGACG	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAAACCTAC	CAACAACCTCT	2340
	AGCCAAGGTT	TTTGTGGTAC	TATGGGATCA	GGAATGAAAA	ATGGAGGGCA	GGAAACCATT	2400
15	GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCCTGGACT	CCTGCAGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACTTACTCG	2520
	GAGTGGCACA	GTTTTACTCA	ACCCCGTCTC	GGTGAAAAAT	TGCATCGATG	TAATCAGAAT	2580
	GAAGACCGCA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGATCT	2640
	CCAGCTGGTT	CTGTGGGCTG	CTGCAGTGAA	AAGCAGGAAG	AAGATGGCCT	TGACTTTTFA	2700
20	AATAATTTGG	AACCCAAATT	TATTACATTA	GCAGAAGCAT	GCACAAAGAG	ATAATGTCAC	2760
	AGTGCTACAA	TTAGGTCITT	GTGAGACATT	CTGAGGTTT	CCAAAAATAA	TATTGTAAAG	2820
	TTCAATTTCA	ACATGTATGT	ATATGATGAT	TTTTTTCTCA	ATTTTGAATT	ATGCTACTCA	2880
	CCAATTTATA	TTTTTAAAGC	CAGTTGTTGC	TTATCTTTTC	CAAAAAGTGA	AAAATGTTAA	2940
25	AACAGACAAC	TGGTAAATCT	CAAACTCCAG	CACCTGGAAT	AAGGTCTCTA	AAGCATCTGC	3000
	TCTTTTTTTT	TTTTACGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAACA	3060
	ATAGCTAAGT	TATGCTAATA	TCACATTATT	ATGTATTAC	TTTAAGTGAT	AGTTTAAAAA	3120
	ATAAACAAGA	AATATTGAGT	ATCACTATGT	GAAGAAAGTT	TTGGAAAAGA	AACAATGAAG	3180
	ACTGAATTAA	ATTAAAAATG	TTGCAGCTCA	TAAAGCAATTG	GGACTCACCC	CTACTGCACT	3240
30	ACCAAAATTCA	TTTGACTTTG	GAGGCAAAAT	GTGTTGAAGT	GCCCTATGAA	GTAGCAATTT	3300
	TCTATAGGAA	TATAGTTGGA	AATAAATGTG	TGTGTGTATA	TTATTATTAA	TCAATGCAAT	3360
	ATTTAAAATG	AAATGAGAAC	AAAGAGGAAA	ATGGTAAAAA	CTTGAAATGA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCCTAGG	CCTGGGCTCT	TAAATGCTGC	3480
	ATTATAACTG	AGTCTATGAG	GAAATAGTTC	CTGTCCAATT	TGTGTAATTT	GTTTAAAATT	3540
35	GFAAATAAAT	TAAACTTTTC	TGGTTTCTGT	GGGAAGGAAA	TAGGGAATCC	AATGGAACAG	3600
	TAGCTTTGCT	TTGCAGTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATACTCGC	TGCAGCTGGG	GTTCCCTGCT	TTTTGGTAGC	AAGGGTCCAG	AGATGAGGTG	3720
	TTTTTTTCGG	GGAGCTAATA	ACAAAAACAT	TTTAAAACCT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATTGCT	GTTTCTATTCT	TCTCTTATAG	TGACCAACAT	CTTTTTAATT	TAGATCCAAA	3840
40	TAACCATGTC	CTCCTAGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
	GCACCCTGGG	GAGATTGATT	GTCTTTAAAC	CTAAGCCCCA	CAAACTTGAC	ACCTGATCAG	3960
	GTCTGGGAGC	TACAAAATTT	CATTTTCTCT	CTCACTGCCC	TTCTTCTGAG	TGGCATTGGC	4020
	CTGAATCAAG	GAAAGCCAGG	CCTTGTGGGC	CCCCTTCTTT	CGGCTTTCTG	CTAAAGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAAGTGACTC	CAGGTTTTC	ACCATCCTTC	AGCGTGAATT	4140
45	AAATTTTAAT	CAGTTTGCTT	TCTCCAGAGA	AAATTTAAAA	TAATAGAAGA	AATAGAAATT	4200
	TTGAATGTAT	AAAAGAAAAA	GATCAAGTTG	TCATTTTAGA	ACAGAGGGAA	CTTTGGGAGA	4260
	AAGCAGCCCA	AGTAGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCTTCAA	4320
	GGGCAAGGAG	AGGCCACAAG	GAATATGGGT	GGAGTAAAAA	GCAACATCGT	CTGCTTCATA	4380
	CTTTTTCCTA	GGCTTGGCAC	TGCCCTTTTCC	TTTCTCAGGC	CAATGGCAAC	TGCCATTTGA	4440
50	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTTGGA	GTGAGAAATC	4500
	AAGGAGACAG	AGCTGACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACTG	4560
	GTTGTGCAGA	ACAAACAAGG	CATTATGAGG	AATTGTTGTA	TTCTTCTGCT	AGCCCTCCTT	4620
	CTGGGCACTA	AGAAGGTCTA	TGAATTAAT	GCCTATCTAA	AATTCTGATT	TATTCCTACA	4680
	TTTTCTGTTT	TCTAATTTGA	CCCTAAAATC	TATGTGTTT	AGACTTAGAC	TTTTTATTGC	4740
55	CCCCCCCCC	TTTTTTTTTG	AGACGGAGTC	TCGCTCTGAC	GCACAGGCTG	GAGTGCAGTG	4800
	GCTCCGATCT	CTGCTCACTG	AAAGCTCCGC	CTCCCGGGTT	CATGCCATTCT	TCCTGCCTCA	4860
	GCCTCCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCACG	CCCGGCTAAT	TTTTTGTATT	4920
	TTAATAGAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTCGATCTCC	TGACCTCGTG	4980
	ATCCGCCTGC	CTCGGCCTCC	CAAAGTGCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040
60	CTGTGTTTCC	GTTTAAAGTC	GTCTTCTTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAAGT	5100
	TGATCATACG	AATTGGATCA	ATCTTGAAAT	ACTCAACCAA	AAGACAGTCG	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGGCA	CAAAATATTG	GTCTGAGAAT	GGAATTCTCT	GTAAGCCTAG	5220
	TTGCTGAAAT	TTCTGCTGT	AACCAGAAGC	CAGTTTATATC	TAACGGCTAC	TGAAACACCC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACTCACCG	ATCAAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
65	CTAGTGCCGA	TAAACTTTCT	CAAAGAGCAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
	TAACCATCTC	TTTGTTCCTT	GAACATGCTG	AAAACACCT	GGTCTGCATG	TATGCCCGAA	5460
	TTTGTAATTC	TTTTCTCTCA	AATGAAATT	TAAATTTAGG	GATTCATTTT	TATATTTTCA	5520
	CATATGTAGT	ATTATTATTT	CCTTATATGT	GTAAGGTGAA	ATTTATGGTA	TTTGAGTGTG	5580
	CAAGAAAATA	TATTTTTTAA	GCTTTCATTT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640
70	TGTAAATATA	CAGAATGTTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAAATGCAGT	5700
	GGGGTTTGTT	TTGCAATGTT	TTAAACAGAG	TTTGTAGTAT	GCTATTAAAA	GAAGTTACTT	5760
	TGCTTTTAAA	GAAACTTGGC	TGCTTAAAA	AAGCAAAAAT	TGGATGCATA	AAGTAAATTT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAACAAT	ATTAACCTTG	TTTCTTGTTT	TTGCTGTATT	5880
	TAGAGATTAA	ATAATTCTAA	GATGATCACT	TTGCAAAAAT	ATGCTTATGG	CTGGCATGGA	5940
75	AATAGAAAATA	CTCAATTATG	TCTTTGTTGT	ATTAATGGGG	AATATTTTGG	ACAATGTTTC	6000
	ATTATCAAAAT	TGTCGACATC	ATTAATATAT	ATTGTAATGT	TGGGAAGAGA	TCATATTTT	6060
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAATT	TTGATCGGGT	6120
	ATTAAAAGTA	TTAGAAGGTG	GTTATAATTG	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180
	AGGGGTTTTA	CCTTGAGGAC	CAGTGTATGC	AAGGGAAAAAC	ATGAGTTAAA	AAGAAAAAGCA	6240
80	GGCAATATTG	CAGTCTTGAT	TCTGCCACTT	ACAGGATAGA	TAATGCCTGA	ACTTTAATGA	6300
	CAAGATGATC	CAACCATAAA	GGTGCTCTGT	GCTTCACAGT	GAATCTTTTC	CCCATGCAGG	6360
	AGTGTGCTCC	CCTACAAACG	TTAAGACTGA	TCATTTCAAA	AATCTATTAG	CTATATCAAA	6420
	AGCCTTACAT	TTTAATATAG	GTTGAACCAA	AATTTCAATT	CCAGTAACTT	CTATTGTAAC	6480
	CATTATTTTT	GTGTATGTCT	TCAAGAATGT	TCATTGGATT	TTTGTTTGTA	ATAGTAAAAAT	6540
85	ACCGGATACA	TTTCACGTGT	CCTTCAGTAT	TGATTTGGTT	GAATATTGGG	TCATAATGGT	6600
	TGAGAAGCAT	GGACACTAGA	GCCAGAATGC	TTGGATATGA	ATCCTGGATC	TGTCACCTAC	6660
	TTCTGTGTGA	CCTTTGAAAG	GCTACTTATT	TCCTCTCTTA	GCTTTCTCAT	TAAAATCAAT	6720
	GAACAAATGCC	AGCCTCATGG	GGTTGTTGAA	TGATTAAATT	AGTTAATATA	CCTAAAGTAC	6780



ATAGAACACT GCCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840  
 GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTTT AGGGAAATAA AGTTTGTGCA 6900  
 TATATATAAT CCCGAAACAT G

5 Seq ID NO: 32 Protein sequence:  
 Protein Accession #: NP\_001932.1

10 1 11 21 31 41 51  
 | | | | | |  
 MAAAGPERSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60  
 ADLIRSSDPD FRVLNDGSVY TARAVALSDK KRSFTIWLSD KRKQTQKEVT VLLEHQKKVS 120  
 KTRHRTRETVL RRAKRRWAPI PCSMQENSLG PFPLFLQQVE SDAAQNYTVF YSISGRGVDK 180  
 EPLNLFYIER DTGNLFCRTP VDREYDVFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240  
 15 PVFTEAIYNF EVLESSRPGT TVGVVCATDR DEPDTHMTRL KYSILQQTTPR SPGLFSVHPS 300  
 TGVITTVSHY LDREVVDKYS LIMKVQMDMG QFFGLIGTST CIITVTDSDND NAPTFRQNAV 360  
 EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKGNE NGHFKISTDK ETNEGVLSVV 420  
 KPLNYEENRQ VNLEIGVNNE APFARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI 480  
 KENLAVGSKI NGYKAYDPEN RNGNGLRYKK LHDPKGWITI DEISGSIITS KILDREVETP 540  
 20 KNELYNITVL AIDKDDRSCT GTLAVNIEDV NDNPPPEILQE YVVICKPKMG YTDILAVDPD 600  
 EPVHGAPFYF SLPNTSPEIS RLWSLTQVND TAARLSYQKN AGFQEYTIPI TVKDRAGQAA 660  
 TKLLRVNLCE CTHPTQCRAT SRSTGVILGK FAILAILLGI ALLFSVLLTL VCGVFGATKG 720  
 KRFPEDLAQQ NLIIISNTEAP GDDRVCSANG WMTQTNNSS QGFCGTMGSG MKNGGQETIE 780  
 MMKGGNQTL ESCRAGHHHT LDSCRGGHTE VDNCRYTYSE WHSFTQPRLG EKLHRCNQNE 840  
 25 DRMPSQDYVL TYNYEGRGSP AGSVGCCSEK QEEDGLDFLN NLEPKFITLA EACTKR

30 Seq ID NO: 33 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 64-2583

35 1 11 21 31 41 51  
 | | | | | |  
 GGCAGGTCTC GCTCTCGGCA CCCTCCCGGC GCCCGCGTTC TCCTGGCCCT GCCCGGCATC 60  
 CCGATGGCCG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120  
 35 CTGACCCCTG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180  
 CCTTCTAAAC TAGAGGCAGA CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240  
 TCTGCAGACC TCATCCGGTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGGTCACTG 300  
 TACACAGCCA GGGCTGTTGC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360  
 40 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420  
 TCGAAGACAA GACACACTAG AGAAACTGTT CTCAGCGGTG CCAAGAGGAG ATGGGCACCT 480  
 ATTCCTTGCT CATTCGAA GAATTCCTTG GCCCTTTCC CATTGTTTCT TCAACAAGTT 540  
 GAATCTGATG CAGCACAGAA CTATACTGTC TTCTACTCAA TAAGTGGACG TGGAGTTGAT 600  
 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG 660  
 45 CCTGTGGATC GTGAAGAATA TGATGTTTTT GATTTGATTG CTTATGCGTC AACTGCAGAT 720  
 GGATATTGAG CAGATCTGCC CCTCCCACTA CCCATCAGGG TAGAGGATGA AAATGACAAC 780  
 CACCCTGTTT TCACAGAAGC AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840  
 ACTCAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGCGC 900  
 CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACCTG GGCTCTTTTC TGTGCATCCC 960  
 50 AGCACAGGCG TAATCACCAC AGTCTCTCAT TATTGGACA GAGAGGTTGT AGACAAGTAC 1020  
 TCATTGATAA TGAAAGTACA AGACATGGAT GGCCAGTTT TTGGATTGAT AGGCACATCA 1080  
 ACTTGATACA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACTTTCAG ACAAATGCT 1140  
 TATGAAGCAT TTGTAGAGGA AATGCATTG AATGTGGAAA TCTTACGAAT ACCTATAGAA 1200  
 GATAAGGATT TAATAACAC TGCCAATTGG AGACTCAATT TTACCATTTT AAAGGGAAAT 1260  
 55 GAAAAATGAC ATTTCAAAAT CAGCACAGAC AAAGAACTA ATGAAGGTGT TCTTTCTGTT 1320  
 GTAAAGCCAC TGAATTATGA AGAAAACCGT CAAGTGAACC TGGAAATTGG AGTAAACAAT 1380  
 GAAGCGCCAT TTGCTAGAGA TATTTCCAGA GTGACAGCCT TGAACAGAGC CTTGGTTACA 1440  
 GTTCATGTGA GGGATCTGGA TGAGGGGCCT GAATGCACTC CTGCAGCCCA ATATGTGCCG 1500  
 ATTAAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA 1560  
 60 AATAGAAATG GCAATGGTTT AAGGTACAAA AAATGTGATG ATCCTAAAGG TTGGATCACC 1620  
 ATTGATGAAA TTTGAGGTC AATCATAACT TCCAAAATCC TGGATAGGGA GGTGAAACT 1680  
 CCCAAAATG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAAGATGA TAGATCATGT 1740  
 ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAAATGATA ATCCACCAGA AATACTTCAA 1800  
 GAATATGTAG TCATTTGCAA ACCAAAAATG GGGTATACCG ACATTTTAGC TGTGATCCCT 1860  
 65 GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC 1920  
 AGTAGACTGT GGAGCCTCAC CAAAGTTAAT GATACAGCTG CCCGTCTTTC ATATCAGAAA 1980  
 AATGCTGGAT TTCAAGAATA TACCATTCCCT ATTACTGTAA AAGACAGGGC CGGCCAAGCT 2040  
 GCAACAAAAT TATTGAGAGT TAATCTGTGT GAATGTACTC ATCCAACCTA GTGTCGTGCG 2100  
 ACTTCAAGGA GTACAGGAGT AATACTTGGA AAATGGGCAA TCCTTGCAAT ATTACTGGGT 2160  
 70 ATAGCACTGC TCTTTTCTGT ATTGCTAACT TTAGTATGTG GAGTTTTTGG TGCAACTAAA 2220  
 GGGAAACGTT TTCCTGAAGA TTTAGCACAG CAAAACCTAA TTATATCAAA CACAGAAGCA 2280  
 CCTGGAGACG ATAGAGTGTG CTCTGCCAAT GGATTATGA CCCAACTAC CAACAACTCT 2340  
 AGCCAAGGTT TTTGTGGTAC TATGGGATCA GGAATGAAAA ATGGAGGGCA GGAAACCATT 2400  
 GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCCTGCC GGGGGGCTGG GCATCATCAT 2460  
 75 ACCCTGGACT CCTGCAGGGG AGGACACACG GAGGTGGACA ACTGCAGATA CACTTACTCG 2520  
 GAGTGGCACA GTTTTACTCA ACCCCGTCTC GGTGAAGAAT CCATTAGAGG ACACACTGGT 2580  
 TAAAAATTAA ACATAAAAGA AATTGCATCG ATGTAATCAG AATGAAGACC GCATGCCATC 2640  
 CCAAGATTAT GTCCTCACTT ATAACATGTA GGGAAGAGGA TCTCCAGCTG GTTCTGTGGG 2700  
 CTGCTGCAGT GAAAAGCAGG AAGAAGATGG CTTGACTTTT TTAAATAATT TGGAACCCAA 2760  
 80 ATTTATTACA TTAGCAGAAG CATGCACAAA GAGATAATGT CACAGTGCTA CAATTAGGTC 2820  
 TTTGTCAGAC ATTCTGGAGG TTTCCAAAAA TAATATTGTA AAGTTCAATT TCAACATGTA 2880  
 TGTATATGAT GATTTTTTTC TCAATTTTGA ATTAGTGTAC TCACCAATTT ATATTTTAA 2940  
 AGCCAGTTGT TGCTTATCTT TTCCAAAAAG TGAAAAATGT TAAACAGAC AACTGGTAAA 3000  
 TCTCAAACTC CAGCACTGGA ATTAAGGTCT CTAAAGCATC TGCTCTTTT TTTTTTTACG 3060  
 85 GATATTTTAG TAATAAATAT GCTGATAAAA TATTAGTCCA ACAATAGCTA AGTTATGCTA 3120  
 ATATCACATT ATTATGTATT CACTTTAAGT GATAGTTTAA AAAATAAACA AGAAATATTG 3180  
 AGTATCACTA TGTGAAGAAA GTTTTGGAAA AGAAACAATG AAGACTGAAT TAAATTAATA 3240  
 ATGTTGCAGC TCATAAAGAA TTGGGACTCA CCCCTACTGC ACTACCAAT TCATTTGACT 3300



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60

```

TTGGAGGCAA AATGTGTTGA AGTGCCCTAT GAAGTAGCAA TTTTCTATAG GAATATAGTT 3360
GGAAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTTAAA ATGAAATGAG 3420
AACAAAAGAG AAAATGGTAA AAACCTTGAAA TGAGGCTGGG GTATAGTTTG TCCTACAATA 3480
GAAAAAAGAG AGAGCTTCCT AGGCCTGGGC TCTTAAATGC TGCATTATAA CTGAGTCTAT 3540
GAGGAAATAG TTCCTGTCCA ATTTGTGTAA TTTGTTTAAA ATTGTAAATA AATTAAACTT 3600
TTCCTGGTTT TGTGGGAAGG AAATAGGGAA TCCAATGGAA CAGTAGCTTT GCTTTGCAGT 3660
CTGTTTCAAG ATTTCTGCAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT 3720
GGGGTTCCTT GCTTTTGTGT AGCAAGGGTC CAGAGATGAG GTGTTTTTTT CGGGGAGCTA 3780
ATAACAAAAA CATTTTAAAA CTTACCTTTA CTGAAGTTAA ATCCTCTATT GCTGTTTCTA 3840
TTCTCTCTTA TAGTGACCAA CATCTTTTTA ATTTAGATCC AAATAACCAT GTCCTCCTAG 3900
AGTTTAGAGG CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAAGCACCTT GGGGAGATTG 3960
ATTGTCCTTA AACCTAAGCC CCACAAACTT GACACCTGAT CAGGTCTGGG AGCTACAAAA 4020
TTTCATTTTT CTCTCACTG CCCTTCTTCT GAGTGGCATT GGCCTGAATC AAGGAAAGCC 4080
AGGCCTTGTG GGCCCCCTTC TTTGGGCTTT CTGCTAAAGC AACACCTCCA GCAGAGATTG 4140
CCTTAAGTGA CTCCAGTTT TCCACCATTC CTGAGCGTGA ATTAATTTTT AATCAGTTTG 4200
CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTTGAATG TATAAAAGAA 4260
AAAGATCAAG TTGTCATTTT AGAACAGAGG GAACTTTGGG AGAAAGCAGC CCAAGTAGGT 4320
TATTTGTACA GTCAGAGGGC AACAGGAAGA TGCAGGCCCT CAAGGGCAAG GAGAGGCCAC 4380
AAGGAATATG GGTGGGAGTA AAAGCAACAT CGTCTGCTTC ATACTTTTTT CTAGGCTTGG 4440
CACTGCCTTT TCCTTTCTCA GGCCAAATGG AACTGCCATT TGAGTCCGGT GAGGGATCAG 4500
CCAACCTCTT CTCTATGGCT CACCTTATTT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC 4560
TGCATGATGA GTCTGAAGGC ATTTGCAGGA TGAGCCTGAA CTGGTTGTGC AGAACAAACA 4620
AGGCATTCTT GGGAAATTGT GTATTCCTTC TGCAGCCCTC CTTCTGGGCA CTAAGAAGGT 4680
CTATGAATTA AATGCCTATC TAAAATTCTG ATTTATTCCT ACATTTTCTG TTTTCTAATT 4740
TGACCTTAAA ATCTATGTGT TTTAGACTTA GACTTTTTAT TGCCCCCCCC CCCTTTTTTT 4800
TTGAGACGGA GTCTCGCTCT GACGCACAGG CTGGAGTGCA GTGGCTCCGA TCTCTGCTCA 4860
CTGAAAGCTC CGCCTCCCGG GTTCATGCCG TTCTCCTGCC TCAGCCTCCT GAGTAGCTGG 4920
GACTACAGGC GCCCCACCAC ACGCCCGGCT AATTTTTTGT ATTTTAAATA GAGACGGGGT 4980
TTCACTGTGT TAGCCAGGAT GGTCTCGATC TCCTGACCTC GTGATCCGCC TGCCTCGGCC 5040
TCCCAAAGTG CTGGGATTAC AGGCATGACC CACCGCTCCC GGCCTTGTTT TCCGTTTAAA 5100
GTCTCTTCTT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTTGATCAT ACGAATTGGA 5160
TCAATCTTGA AATACTCAAC CAAAAGACAG TCGAGAAGCC AGGGGGAGAA AGAACTCAGG 5220
GCACAAAATA TTGGTCTGAG AATGGAATCT TCTGTAAGCC TAGTTGCTGA AATTTCTGCT 5280
TGTAACCAGA AGCCAGTTTT ATCTAACGGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340
CCCACTCACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTGCCG ATAACTTTC 5400
TCAAAGAGCA ACCAGTATCA CTTCCCTGTT TATAAACCT CTAACCATCT CTTTGTCTCT 5460
TGAACTATGCT GAAAACCACC TGGTCTGCAT GTATGCCCGA ATTTGTAATT CTTTTCTCTC 5520
AAATGAAAAT TTAATTTTAG GGATTCATTT CTATATTTTC ACATATGTAG TATTATTATT 5580
TCCTTATATG TGTAAGGTGA AATTTATGGT ATTTGAGTGT GCAAGAAAAT ATATTTTAA 5640
AGCTTTCATT TTTCCCCAG TGAATGATTT AGAATTTTTT ATGTAAATAT ACAGAAATGT 5700
TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAAATGCAG TGGGGTTTGT TTTGCAATGT 5760
TTTAAACAGA GTTTTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAACTTGG 5820
CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAAGTAATAT TTACAGATGT GGGGAGATGT 5880
AATAAAACAA TATTAACCTG GCTGCTTAAA ATAAGCAAAA ATTGGATGCA TAAAGTAATA 5940
TTTACAGATG TGGGGAGATG TAATAAAACA ATATTAACCT GGTTTCTTGT TTTTGCTGTA 6000
TTTAGAGATT AAATAATTCT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCTGGCATG 6060
GAAATAGAAA TACTCAATTA TGTCTTTGTT GTATTAATGG GGAATATTTT GGACAATGTT 6120
TCATTATCAA ATTGTGACA TCATTAATAT ATATTGTAAT GTTGGGAAGA GATCACTATT 6180
TTGAAGCACA GCTTTACAGA TGAGTATCTA TGATACATAT GTATAATAAA TTTTGATCGG 6240
GTATTTAAAAG TATTAGAAGG TGGTTATAAT TGCAAGATAT TCCATGAATA GTACACTGAC 6300
ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAA ACATGAGTTA AAAAGAAAAG 6360
CAGGCAATAT TGCAGTCTTG ATTCTGCCAC TTACAGGATA GATAATGCCT GAACCTTAAT 6420
GACAAGATGA TCCAACCATA AAGGTGCTCT TGCTTTCACA GTGAATCTTT TCCCCATGCA 6480
GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTCA AAAATCTATT AGCTATATCA 6540
AAAGCCTTAC ATTTTAATAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600
ACCATTATTT TTGTGTATGT CTTCAAGAAT GTTCATTGGA TTTTGTGTTG TAATAGTAAA 6660
ATACCGGATA CATTTACAGT GTCCTTCAGT ATTGATTTGG TTGAATATTG GGTCAATATG 6720
GTTGAGAAGC ATGGACACTA GAGCCAGAA TCTTGGATAT GAATCCTGGA TCTGTCACTT 6780
ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCT TAGCTTTCTC ATTAATAATCA 6840
ATGAACAATG CCAGCCTCAT GGGGTTGTTG AATGATTAAA TTAGTTAATA TACCTAAAGT 6900
ACATAGAACA CTGCCTGCAC ATAGTAAAAG AATTATAAGT GTGAGGTAGT TGGTAAAATT 6960
ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAAT AAAGTTTGTG 7020
CATATATATA ATCCCGAAAC ATG

```

Seq ID NO: 34 Protein sequence:  
Protein Accession #: NP\_077741.1

70  
75  
80  
85

```

1      11      21      31      41      51
|      |      |      |      |      |
MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60
ADLIRSSDPD FRVLNDGSVY TARAVALSDK KRSFTIWLSD KRKQTQKEVT VLEHQKKVS 120
KTRHRTRETVL RRAKRRWAPI PCSMQENSLG PFPLFLQQVE SDAAQNYTVF YSISGRGVDK 180
EPLNLFYIER DTGNLFCTRTP VDREEYDVFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240
PVFTEAIYNF EVLESSRPGT TVGVVCATDR DEPDTHMTRL KYSILQQTPR SPGLFSVHPS 300
TGVITTVSHY LDREVVDKYS LIMKVQDMDG QFFGLIGTST CIITVTDSND NAPTFRQNAV 360
EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKGNE NGHFKISTDK ETNEGVLSSV 420
KPLNVEENRQ VNLEIGVNNE APFARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQVYRI 480
KENLAVGSKI NGYKAYDPEN RENGNGLRK KLDHDPKGWITI DEISGSIITS KILDREVETP 540
KNELYNITVL AIDKDDRSCT GTLAVNIEDV NDNPPAILQE YVVICKPKMG YTDILAVDPD 600
EPVHGAPFYF SLPNTSPEIS RLWSLTQVND TAARLSYQKN AGFQEYTIPI TVKDRAGQAA 660
TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720
KRFPEDLAQQ NLIISNTEAP GDDRVCSANG FMTQTTNNS QGFCGTMGSG MKNGGQETIE 780
MMKGGNQTL ESCRAGHHHT LDSCRGGHTE VDNCRYTYSE WHSFTQPRLG EESIRGHTG

```

Seq ID NO: 35 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 146-1273.



	1	11	21	31	41	51	
5	GGGAGTGGGC	GTGGCGGTGC	TGCCCAGGTG	AGCCACCGCT	GCTTCTGCCC	AGACACGGTC	60
	GCCTCCACAT	CCAGGTCTTT	GTGCTCCTCG	CTTGCCCTGTT	CCTTTTCCAC	GCATTTTCCA	120
	GGATAACTGT	GACTCCAGGC	CCGCAATGGA	TGCCCTGCAA	CTAGCAAATT	CGGCTTTTGC	180
	CGTTGATCTG	TTCAAACAAC	TATGTGAAAA	GGAGCCACTG	GGCAATGTCC	TCTTCTCTCC	240
	AATCTGTCTC	TCCACCTCTC	TGTCACCTGC	TCAAGTGGGT	GCTAAAGGTG	ACACTGCAAA	300
10	TGAAATTGGA	CAGGTTCTTC	ATTTTGAAAA	TGTCAAAAGAT	ATACCCCTTG	GATTTCAAAC	360
	AGTAACATCG	GATGTAAACA	AACTTAGTTC	CTTTTACTCA	CTGAAACTAA	TCAAGCGGCT	420
	CTACGTAGAC	AAATCTCTGA	ATCTTTCTAC	AGAGTTCATC	AGCTCTACGA	AGAGACCCTA	480
	TGCAAAGGAA	TTGGAAACTG	TTGACTTCAA	AGATAAATTG	GAAGAAACGA	AAGGTCAGAT	540
	CAACAACCTCA	ATTAAGGATC	TCACAGATGG	CCACTTTGAG	AACATTTTAG	CTGACAACAG	600
15	TGTGAACGAC	CAGACCAAAA	TCCTTGTGGT	TAATGCTGCC	TACTTTGTTG	GCAAGTGGAT	660
	GAAGAAATTT	CCTGAATCAG	AAACAAAAGA	ATGTCCTTTC	AGACTCAACA	AGACAGACAC	720
	CAAACCAAGTG	CAGATGATGA	ACATGGAGGC	CACGTTCTGT	ATGGGAAACA	TTGACAGTAT	780
	CAATTGTAAG	ATCATAGAGC	TTCTTTTCA	AAATAAGCAT	CTCAGCATGT	TCATCCTACT	840
	ACCCAAGGAT	GTGGAGGATG	AGTCCACAGG	CTTGGAGAAG	ATTGAAAAAC	AACTCAACTC	900
20	AGAGTCACATG	TCACAGTGGA	CTAATCCCAG	CACCATGGCC	AATGCCAAGG	TCAAACCTCTC	960
	CATTCCAAAA	TTTAAGGTGG	AAAAGATGAT	TGATCCCAAG	GCTTGTCTGG	AAAATCTAGG	1020
	GCTGAAACAT	ATCTTCAGTG	AAGACACATC	TGATTTCTCT	GGAATGTCAG	AGACCAAGGG	1080
	AGTGGCCCTA	TCTATCTTTT	TCCACAAGT	TGCTTTAGAA	ATAACTGAAG	ATGGTGGGGA	1140
	TTCCATAGAG	GTGCCAGGAG	CACGGATCCT	GCAGCACAA	GATGAATTGA	ATGCTGACCA	1200
25	TCCCTTTTAT	TACATCATCA	GGCACAACAA	AACTCGAAAC	ATCATTTTCT	TTGGCAAATT	1260
	CTGTTCTCCT	TAAGTGGCAT	AGCCCATGTT	AAGTCCCTCC	TGACTTTTCT	GTGGATGCCG	1320
	ATTTCTGTAA	ACTCTGCATC	CAGAGATTCA	TTTTCTAGAT	ACAATAAATT	GCTAATGTTG	1380
	CTGGATCAGG	AAGCCGCCAG	TACTTGTGAT	ATGTAGCCTT	CACACAGATA	GACCTTTTTT	1440
	TTTTTCCAAT	TCTATCTTTT	GTTTCCTTTT	TTCCATAAAG	ACAATGACAT	ACGCTTTTAA	1500
30	TGAAAAGGAA	TCACGTTAGA	GGAAAAATAT	TTATTCATTA	TTTGTCAAAT	TGTCCGGGGT	1560
	AGTTGGCAGA	AATACAGTCT	TCCACAAAGA	AAATTCCTAT	AAGGAAGATT	TGGAAGCTCT	1620
	TCTTCCCAGC	ACTATGCTTT	CCTTCTTTGG	GATAGAGAAT	GTTCCAGACA	TTCTCGCTTC	1680
	CCTGAAAGAC	TGAAGAAAGT	GTAGTGCATG	GGACCCACGA	AACTGCCCTG	GCTCCAGTGA	1740
	AACTTGGGCA	CATGCTCAGG	CTACTATAGG	TCCAGAAGTC	CTTATGTTAA	GCCCTGGCAG	1800
35	GCAGGTGTTT	ATTAAATTC	TGAATTTTGG	GGATTTTCAA	AAGATAATAT	TTTACATACA	1860
	CTGTATGTTA	TAGAACTTCA	TGGATCAGAT	CTGGGGCAGC	AACCTATAAA	TCAACACCTT	1920
	AATATGCTGC	AACAAAATGT	AGAATATTCA	GACAAAATGG	ATACATAAAG	ACTAAGTAGC	1980
	CCATAAGGGG	TCAAATTTTG	CTGCCAAATG	CGTATGCCAC	CAACTTACAA	AAACACTTCG	2040
	TTTCGACAGC	TTTTCAGATT	GTGGAATGTT	GGATAAGGAA	TTATAGACCT	CTAGTAGCTG	2100
40	AAATGCAAGA	CCCCAAGAGG	AAGTTCAGAT	CTTAATATAA	ATTCACCTTC	ATTTTGTGATA	2160
	GCTGTCCCAT	CTGGTCATGT	GGTTGGCACT	AGACTGGTGG	CAGGGGCTTC	TAGCTGACTC	2220
	GCACAGGGAT	TCTCACAATA	GCCGATATCA	GAATTTGTGT	TGAAGGAACT	TGTCTCTTCA	2280
	TCTAATATGA	TAGCGGGAAA	AGGAGAGGAA	ACTACTGCCT	TTAGAAAATA	TAAGTAAAGT	2340
	GATTAAAGTG	CTCACGTTAC	CTTGACACAT	AGTTTTTCAG	TCTATGGGTT	TAGTTACTTT	2400
45	AGATGGCAAG	CATGTAACCT	ATATTAATAG	TAATTTGTAA	AGTTGGGTGG	ATAAGCTATC	2460
	CCTGTTGCCG	GTTCATGGAT	TACTTCTCTA	TAAAAAATAT	ATATTTACCA	AAAAATTTTG	2520
	TGACATTCCT	TCTCCCATCT	CTTCCTTGAC	ATGCATTGTA	AATAGGTTCT	TCTTGTCTCTG	2580
	AGATTCAATA	TTGAATTTCT	CCTATGCTAT	TGACAATAAA	ATATTATTGA	ACTACC	

Seq ID NO: 36 Protein sequence:  
Protein Accession #: NP\_002630.1

	1	11	21	31	41	51	
55	MDALQLANSA	FAVDLQKQLC	EKEPLGNVLF	SPICLSTSL	LAQVGAKGDT	ANEIGQVLHF	60
	ENVKDIPFGF	QTVTSDVNKL	SSFYSLKLIK	RLYVDKSLNL	STEFISSTKR	PYAKELETVD	120
	FKDKLEETKG	QINNSIKDLT	DGHFENILAD	NSVNDQTKIL	VVNAAYFVGK	WMKKFPESET	180
	KECPFRNLKT	DTKPVQMMNM	EATFCMGNID	SINCKIIELP	FQNKHLMSFI	LLPKDVEDES	240
	TGLEKIEKQL	NSESLSQWTN	PSTMANAKVK	LSIPKFKVEK	MIDPKACLEN	LGLKHIFSED	300
60	TSDFSGMSET	KGVALSNVIH	KVCLEITEDG	GDSIEVPGAR	ILQHKDELNA	DHPFIYIIRH	360
	NKTRNIIFFG	KFCSP					

Seq ID NO: 37 DNA sequence  
Nucleic Acid Accession #: NM\_0168583  
Coding sequence: 72-842

	1	11	21	31	41	51	
70	GGAGTGGGGG	AGAGAGAGGA	GACCAGGACA	GCTGCTGAGA	CCTCTAAGAA	GTCCAGATAC	60
	TAAGAGCAAA	GATGTTTCAA	ACTGGGGGCC	TCATTTGCTT	CTACGGGCTG	TTAGCCCAGA	120
	CCATGGCCCA	GTTTGGAGGC	CTGCCCGTGC	CCCTGGACCA	GACCCTGCCC	TTGAATGTGA	180
	ATCCAGCCCT	GCCCTTGAGT	CCCACAGGTC	TTGCAGGAAG	CTTGACAAAT	GCCCTCAGCA	240
	ATGGCCTGCT	GTCTGGGGGC	CTGTTGGGCA	TTCTGGAAAA	CCTTCCGCTC	CTGGACATCC	300
75	TGAAGCCTGG	AGGAGGTACT	TCTGGTGGCC	TCCTTGGGGG	ACTGCTTGGA	AAAGTGACGT	360
	CAGTGATTCC	TGGCCTGAAC	AACATCATTG	ACATAAAGGT	CACTGACCCC	CAGCTGCTGG	420
	AACTTGGCCT	TGTGCAGAGC	CCTGATGGCC	ACCGTCTCTA	TGTCACCATC	CCTCTCGGCA	480
	TAAAGCTCCA	AGTGAATACG	CCCCTGGTGC	GTGCAAGTCT	GTTGAGGCTG	GCTGTGAAGC	540
	TGGACATCAC	TGACAGAAATC	TTAGCTGTGA	TGACATAAGCA	GGAGAGGATC	CACCTGGTCC	600
80	TTGGTGACTG	CACCCATTCC	CCTGGAAGCC	TGCAAATTTT	TCTGCTTGAT	GGACTTGGCC	660
	CCCTCCCCAT	TCAAGGTCTT	CTGGACAGCC	TCACAGGGAT	CTTGAATAAA	GTCCTGCCTG	720
	AGTTGGTTCA	GGGCAACGTG	TGCCCTCTGG	TCAATGAGGT	TCTCAGAGGC	TTGGACATCA	780
	CCCTGGTGCA	TGACATTGTT	AACATGCTGA	TCCACGGACT	ACAGTTTGTC	ATCAAGGTCT	840
	AAGCCTTCCA	GGAAGGGGCT	GGCCTCTGCT	GAGCTGCTTC	CCAGTGCTCA	CAGATGGCTG	900
85	GCCCATTGTC	TGGAAGATGA	CACAGTTGCC	TTCTCTCCGA	GGAACCTGCC	CCCTCTCCTT	960
	TCCCACCAGG	CGTGTGTAA	ATCCCATTGTG	CCTCACCTAA	TAAAAATGGCT	CTTCTTCTGC	1020
	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA				



Seq ID NO: 38 Protein sequence:  
Protein Accession #: NP\_057667

5	1	11	21	31	41	51	
	MFQTGGLIVF	YGLLAQDMAQ	FGGLPVPLDQ	TLPLNVNPAL	PLSPTGLAGS	LTNALSNGLL	60
	SGGLLGILEN	LPLLDILKPG	GGTSGGGLLG	LLGKVTSVIP	GLNNIIDIKV	TDPQLLELGL	120
	VQSPDGHRLY	VTIPLGIKLO	VNTPLVGASL	LRLAVKLDIT	AEILAVRDQK	ERHLVVLGDC	180
10	THSPGSLQIS	LDDGLGPLPI	QGLLDSLTI	LNKVLPELVQ	GNVCPLVNEV	LRGLDITLVH	240
	DIVNMLIHGL	QFVIKV					

Seq ID NO: 39 DNA sequence  
Nucleic Acid Accession #: NM\_004363.1  
Coding sequence: 115-2223

	1	11	21	31	41	51	
20	CTCAGGGCAG	AGGGAGGAAG	GACAGCAGAC	CAGACAGTCA	CAGCAGCCTT	GACAAAACGT	60
	TCCTGGAACT	CAAGCTCTTC	TCCACAGAGG	AGGACAGAGC	AGACAGCAGA	GACCATGGAG	120
	TCTCCCTCGG	CCCCCTCCCA	CAGATGGTGC	ATCCCCCTGGC	AGAGGCTCCT	GCTCACAGCC	180
	TCACCTCTAA	CCTTCTGGAA	CCCGCCACC	ACTGCCAAGC	TCACTATTGA	ATCCACGCCG	240
	TTCAATGTCT	CAGAGGGGAA	GGAGGTGCTT	CTACTTGTCC	ACAATCTGCC	CCAGCATCTT	300
25	TTTGGCTACA	GCTGGTACAA	AGGTGAAAGA	GTGGATGGCA	ACCGTCAAAT	TATAGGATAT	360
	GTAATAGGAA	CTCAACAAGC	TACCCACAGG	CCCGCATACA	GTGGTCGAGA	GATAATATAC	420
	CCCAATGCAT	CCCTGCTGAT	CCAGAACATC	ATCCAGAATG	ACACAGGATT	CTACACCCTA	480
	CACGTCATAA	AGTCAGATCT	TGTGAATGAA	GAAGCAACTG	GCCAGTTCCG	GGTATACCCG	540
	GAGCTGCCCA	AGCCCTCCAT	CTCCAGCAAC	AACTCCAAAC	CCGTGGAGGA	CAAGGATGCT	600
	GTGGCCTTCA	CCTGTGAACC	TGAGACTCAG	GACGCAACCT	ACCTGTGGTG	GGTAAACAAT	660
30	CAGAGCCTCC	CGGTCACTCC	CAGGCTGCAG	CTGTCCAATG	GCAACAGGAC	CCTCACTCTA	720
	TTCAATGTCA	CAAGAAATGA	CACAGCAAGC	TACAAATGTG	AAACCCAGAA	CCCAGTGAGT	780
	GCCAGGCGCA	GTGATTCACT	CATCCTGAAT	GTCTCTATG	GCCCCGATGC	CCCCACCAT	840
	TCCCCTCTAA	ACACATCTTA	CAGATCAGGG	GAAAATCTGA	ACCTCTCCTG	CCACGCAGCC	900
35	TCTAACCAC	CTGCACAGTA	CTCTTGGTTT	GTCAATGGGA	CTTTCAGCA	ATCCACCCAA	960
	GAGCTCTTTA	TCCCCAACAT	CACGTGTGAAT	AATAGTGGAT	CCTATACGTG	CCAAGCCCAT	1020
	AACTCAGACA	CTGGCCTCAA	TAGGACCACA	GTCACGACGA	TCACAGTCTA	TGCAGAGCCA	1080
	CCCCAACCCCT	TCATCACCAG	CAACAACCTCC	AACCCCGTGG	AGGATGAGGA	TGCTGTAGCC	1140
	TTAACCTGTG	AACCTGAGAT	TCAGAACACA	ACCTACCTGT	GGTGGGTAA	TAATCAGAGC	1200
40	CTCCCGGTCA	GTCCCAGGCT	GCAGCTGTCC	AATGACAACA	GGACCCCTAC	TCTACTCAGT	1260
	GTCACAAGGA	ATGATGTAGG	ACCCTATGAG	TGTGGAATCC	AGAACGAATT	AAGTGTGTGAC	1320
	CACAGCGACC	CAGTCATCCT	GATGTCTCTC	TATGGCCAG	ACGACCCAC	CATTTCCCCC	1380
	TCATACACCT	ATTACCGTCC	AGGGGTGAAC	CTCAGCCTCT	CCTGCCATGC	AGCCTCTAAC	1440
	CCACCTGCAC	AGTATTCTTG	GCTGATTGAT	GGGAACATCC	AGCAACACAC	ACAAGAGCTC	1500
45	TTTATCTCCA	ACATCACTGA	GAAGAACAGC	GGACTCTATA	CCTGCCAGGC	CAATAACTCA	1560
	GCCAGTGGCC	ACAGCAGGAC	TACAGTCAAG	ACAATCACAG	TCTCTGCGGA	GCTGCCCCAAG	1620
	CCCTCCATCT	CCAGCAACAA	CTCCAAACCC	GTGGAGGACA	AGGATGCTGT	GGCCTTCACC	1680
	TGTGAACCTG	AGGCTCAGAA	CACAACCTAC	CTGTGGTGGG	TAAATGGTCA	GAGCCTCCCA	1740
	GTCAGTCCCA	GGCTGCAGCT	GTCCAATGGC	AACAGGACCC	TCACTCTATT	CAATGTCAACA	1800
50	AGAAATGACG	CAAGAGCCTA	TGTATGTGGA	ATCCAGAACT	CAGTGAGTGC	AAACCGCAGT	1860
	GACCCAGTCA	CCCTGGATGT	CCTCTATGGG	CCGGACACCC	CCATCATTTT	CCCCCAGAC	1920
	TCGTCTTACC	TTTCGGGAGC	GAACCTCAAC	CTCTCCTGCC	ACTCGGCCTC	TAACCCATCC	1980
	CCGCAGTATT	CTTGGCGTAT	CAATGGGATA	CCGCAGCAAC	ACACACAAGT	TCTCTTTATC	2040
	GCCAAAATCA	CGCCAAATAA	TAACGGGACC	TATGCCCTGT	TTGTCTCTAA	CTTGGCTACT	2100
55	GGCCGCAATA	ATTCCATAGT	CAAGAGCATC	ACAGTCTCTG	CATCTGGAAC	TTCTCCTGGT	2160
	CTCTCAGCTG	GGGCCACTGT	CGGCATCATG	ATTGGAGTGC	TGGTTGGGGT	TGCTCTGATA	2220
	TAGCAGCCCT	GGTGTAGTTT	CTTCATTTCA	GGAAAGACTGA	CAGTTGTTTT	GCTTCTTCCT	2280
	TAAAGCATTT	GCAACAGCTA	CAGTCTAAAA	TTGCTTCTTT	ACCAAGGATA	TTTACAGAAA	2340
	AGACTCTGAC	CAGAGATCGA	GACCATCTTA	GCCAACATCG	TGAAACCCCA	TCTCTACTAA	2400
60	AAATACAAAA	ATGAGCTGGG	CTTGGTGGCG	GCGACCTGTA	GTCCCAGTTA	CTCGGGAGGC	2460
	TGAGGCAGGA	GAATCGCTTG	AACCCGGGAG	GTGGAGATTG	CAGTGAGCCC	AGATCGCACC	2520
	ACTGCACTCC	AGTCTGGCAA	CAGAGCAAGA	CTCCATCTCA	AAAAGAAAAG	AAAAGAAGAC	2580
	TCTGACCTGT	ACTCTTGAAT	ACAAGTTTCT	GATACCACTG	CACTGTCTGA	GAATTTCCAA	2640
	AACTTTAATG	AACTAACTGA	CAGCTTCATG	AAACTGTCCA	CCAAGATCAA	GCAGAGAAAA	2700
65	TAATTAATTT	CATGGGACTA	AATGAACCTA	TGAGGATTGC	TGATTCTTTA	AATGTCTTGT	2760
	TTCCCAGATT	TCAGGAAACT	TTTTTCTTTT	TAAGCTATCC	ACTCTTACAG	CAATTTGATA	2820
	AAATATACTT	TTGTGAACAA	AAATTGAGAC	ATTTACATTT	TCTCCCTATG	TGGTCGCTCC	2880
	AGACTTGGGA	AACTATTTCAT	GAATATTTAT	ATTGTATGGT	AATATAGTTA	TTGCACAAGT	2940
70	TCAATAAAAA	TCTGCTCTTT	GTATAACAGA	AAAA			

Seq ID NO: 40 Protein sequence:  
Protein Accession #: NP\_004354.1

	1	11	21	31	41	51	
75	MESPSAPPHR	WCIPWQRLLL	TASLLTFWNP	PTAKLTIES	TPFNVAEGKE	VLLLVDHNLPO	60
	HLFGYSWKY	ERVDFGNRQII	GYVIGTQQAT	PGPAYSGREI	IYPNASLLIQ	NIIQNDTGFI	120
	TLHVIKSDLV	NEEATGQFRV	YPELPKPSIS	SNNKPFVEDK	DAVAFTCEPE	TQDATYLWVW	180
	NNQSLPVSFR	LQLSNGNRTL	TLFNVTRNDT	ASYKCETQNP	VSARRSDSVI	LNVLVGPDPAP	240
	TISPLNTSYR	SGENLNLSCH	AASNPPAQYS	WVFNQTFQQS	TQELFIPNIT	VNNSGSYTCQ	300
80	AHNSDTGLNR	TTVTITVYA	EPPKPFITSN	NSNPVEDEDA	VALTCEPEIQ	NTTYLWVWNN	360
	QSLPVSFRLQ	LSNDNRTLTL	LSVTRNDVGP	YECGIQNELS	VDHSDPVILN	VLYGPDPTI	420
	SPSYTYRPG	VNLSLSCHAA	SNPPAQYSWL	IDGNIQQHTQ	ELFISNITEK	NSGLYTCQAN	480
	NSASGHSRTT	VKTITVSAEL	PKPSISSNNS	KPVEDKDAVA	FTCEPEAQNT	TYLWVWNGQS	540
	LPVSPRLQLS	NGNRTLTLFN	VTRNDARAYV	CGIQNSVSAN	RSDPVTLDEL	YGPDTPIISP	600
85	PDSSYLSGAN	LNLSCHSASN	PSPQYSWRIN	GIPQOHTQVL	FIAKITPNNN	GTACFVSNL	660
	ATGRNNSIVK	SITVSASGTS	PGLSAGATVG	IMIGVLVGVA	LI		



Seq ID NO: 41 DNA sequence  
Nucleic Acid Accession #: NM\_006952.1  
Coding sequence: 11-793

```

1      11      21      31      41      51
|      |      |      |      |      |
AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTTTCGTTGC TTCCAGGGCC TGCTGATTTT 60
TGGAAATGTG ATTATTGGTT GTTTCGGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG 180
GGCTGCCTGG ATCGGCATAT TTGTGGGCAT CTGCCTCTTC TGCCTGTCTG TTCTAGGCAT 240
TGTAGGCATC ATGAAGTCCA GCAGGAAAAT TCTTCTGGCG TATTTTCATTG TGATGTTTAT 300
AGTATATGCC TTTGAAGTGG CATCTTGTAT CACAGCAGCA ACACAACGAG ACTTTTTCAC 360
ACCCAACCTC TTCCTGAAGC AGATGCTAGA GAGGTACCAA AACAAACAGCC CTCCAAACAA 420
TGATGACCAG TGGAAAAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480
CAATTGCTGT GGCCTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CCTTCCGGAC 540
TGAGAATAAT GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA 600
AGAACCTCTC AACCTGGAGG CTTGTAAACT AGGCGTGCCT GGTTTTTATC ACAATCAGGG 660
CTGCTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CCTGGTTTGG 720
ATTTGCCATT CTCTGCTGGA CTTTTTGGGT TCCTCTGGGT ACCATGTTCT ACTGGAGCAG 780
AATTGAATAT TAAGAA

```

Seq ID NO: 42 Protein sequence:  
Protein Accession #: NP\_008883.1

```

1      11      21      31      41      51
|      |      |      |      |      |
MAKDNSTVRC FQGLLIFGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAW 60
IGIFVGICLF CLSVLGIVGI MKSSRKILLA YFILMFIVYA FEVASCITAA TQRDFFTPNL 120
FLKQMLERYQ NNSPPNDDQ WKNNGVTKTW DRLMLQDNCC GVNGPSDWQK YTSAFRTENN 180
DADYFWRQC CVMNNLKEPL NLEACKLGVP GFYHNQGCYE LISGPMNRHA WGVAVFGFAI 240
LCWTFWVLLG TMFYWSRIEY

```

Seq ID NO: 43 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 83-2605

```

1      11      21      31      41      51
|      |      |      |      |      |
GCCGGACAGA TCTGCGCGTA TCCTGGAGCC GGCCCAGTTG TGAAGTAGGA GAGCTTTGGG 60
ACCTCTGTCC CAAGCAAGAG AGATGAATGG AGAGTATAGA GGCAGAGGAT TTGGACGAGG 120
AAGATTTCAA AGCTGGAAAA GGGGAAGAGG TGGTGGGAAC TTCTCAGGAA AATGGAGAGA 180
AAGAGAACAC AGACCTGATC TGAGTAA AAC CACAGGAAAA CGTACTTCTG AACAAACCCC 240
ACAGTTTTTG CTTTCAACAA AGACCCACA GTCAATGCAG TCAACATTGG ATCGATTCAAT 300
ACCATATAAA GGCTGGAAGC TTTATTTCTC TGAAGTTTAC AGCGATAGCT CTCCTTTGAT 360
TGAGAAGATT CAAGCATTG AAAAAATTTT CACAAGGCAT ATTGATTGTG ATGACAAGGA 420
TGAAATAGAA AGAAAGGGAA GTATTTTGGT AGATTTTAAA GAACTGACAG AAGGTGGTGA 480
AGTAACTAAC TTGATACCAG ATATAGCAAC TGAAGTAAAG GATGCACCTG AGAAAACCTT 540
GGCTTGCAATG GGTTTGGCAA TACATCAGGT GTTAACTAAG GACCTTGAAA GGCATGCAGC 600
TGAGTTACAA GCCCAGGAAG GATTGTCTAA TGATGGAGAA ACAATGGTAA ATGTGCCACA 660
TATTCTATGCA AGGGTGTACA ACTATGAGCC TTTGACACAG CTCAAGAATG TCAGAGCAAA 720
TTACTATGGA AAATACATTG CTCTAAGAGG GACAGTGGTT CGTGTCAAGTA ATATAAAGCC 780
TCTTTGCACC AAGATGGCTT TTCTTTGTGC TGATGTGGA GAAATTCAGA GCTTTCTCT 840
TCCAGATGGA AAATACAGTC TTCCACAAA GTGTCTGTG CCTGTGTGTC GAGGCAGGTC 900
ATTTACTGCT CTCCGAGCT CTCCTCTCAC AGTTACGATG GACTGGCAGT CAATCAAAAT 960
CCAGGAATTG ATGTCTGATG ATCAGAGAGA AGCAGGTCGG ATTCCACGAA CAATAGAATG 1020
TGAGCTTGTT CATGATCTTG TGGATAGCTG TGTCCCGGGA GACACAGTGA CTATTACTGG 1080
AATTGTCAAA GTCTCAAATG CGGAAGAAGG TTCTCGAAAT AAGAATGACA AGTGTATGTT 1140
CCTTTTGTAAT ATTGAAGCAA ATTCTATTAG TAATAGCAAA GGACAGAAAA CAAAGAGTTC 1200
TGAGGATGGG TGTAAGCATG GAATGTTGAT GGAGTTCTCA CTTAAAGACC TTTATGCCAT 1260
CCAAGAGATT CAAGCTGAAG AAAACCTGTT TAAACTCATT GTCAACTCGC TTTGCCCTGT 1320
CATTTTGGT CATGAACCTG TTAAAGCAGG TTTGGCATTA GCACTCTTTG GAGGAAGCCA 1380
GAAATACGCA GATGACAAAA ACAGAATTCC AATTCGGGGA GACCCCAACA TCCTTGTTGT 1440
TGGAGATCCA GGCCTAGGAA AAAGTCAAAT GCTACAGGCA GCGTGCAATG TTGCCCCACG 1500
TGGCGTGAT GTTTGTGTA ACACCACGAC CACCTCTGGT CTGACGGTAA CTC'TTCAAA 1560
AGATAGTTCC TCTGGAGATT TTGCTTTGGA AGCTGGTGCC CTGGTACTTG GTGATCAAGG 1620
TATTTGTGGA ATCGATGAAT TTGATAAGAT GGGGAATCAA CATCAAGCCT TGTGGAAGC 1680
CATGGAGCAG CAAAGTATTA GTCTTGCTAA GGCTGGTGTG GTTTGTAGCC TTCCTGCAAG 1740
AACTTCCATT ATTGCTGCTG CAAATCCAGT TGGAGGACAT TACAATAAAG CCAAAACAGT 1800
TTCTGAGAA TTAATAATGG GAGGTGCACT ACTATCCAGA TTTGATTTGG TCTTTATCCT 1860
GTTAGATACT CCAATGAGC ATCATGATCA CTTACTCTCT GAACATGTGA TTGCAATAAG 1920
AGCTGGAAAG CAGAGAACCA TTAGCAGTGC CACAGTAGCT CGTATGAATA GTCAAGATTC 1980
AAATACTTCC GTACTTGAAG TAGTTTCTGA GAAGCCATTA TCAGAAAGAC TAAAGGTGGT 2040
TCCTGGAGAA ACAATAGATC CCATTCCCCA CCAGCTATTG AGAAAGTACA TTGGCTATGC 2100
TCGGCAGTAT GTGTACCCAA GGCTATCCAT AGAAGCTGCT CGAGTTCTTC AAGATTTTTC 2160
CCTTGAGCTC CGGAAACAGA GCCAGAGGTT AAATAGCTCA CCAATCACTA CCAGGCAGCT 2220
GGAATCTTTG ATTCGTCTGA CAGAGGCACG AGCAAGGTTG GAATTGAGAG AGGAAGCAAC 2280
CAAAGAAGAC GCTGAGGATA TAGTGGAAT TATGAAATAT AGCATGCTAG GAACTTACTC 2340
TGATGAATTT GGAACCTAG ATTTTGAGCG ATCCCAGCAT GGTCTTGGA TGAGCAACAG 2400
GTCAACAGCG AAAAGATTTA TTTCTGCTCT CAACAACGTT GCTGAAAGAA CTTATAATAA 2460
TATATTTCAA TTTTCATCAAC TTCGGCAGAT TGCCAAAGAA CTAAACATTC AGGTTGCTGA 2520
TTTTGAAAAT TTTATTTGAT CACTAAATGA CCAGGGTTAC CTCTTGAAAA AAGGCCCAAA 2580
AGTTTACCAG CTTCAAACTA TGTAAAGGA CTTACCAAG TTAGGGCTC CTGGGTTTAT 2640
TGCAGATTAA AGCCATCTCA GTGAAGATAT GCGTGACGCG ACAGACAGAC AGACACACAC 2700
ACACACACAC ACACACACAC ACACACACAC ACACACAGTC AAATACTGTT CTCTGAAAAA 2760
TGATGTCCCA AAAGTATTAT AATAGGAAAA AAGCATTAAT TATAATAAAC TAATTTAAGA 2820

```



AGTGATAAAG TCTCCAGATG CAGTAGCTCA CACTGTAATC ACAGTGAATC AGGAGGCTGA 2880  
GGTGAGAGGA TTCCTTGAGG CCAGGGTTTCG AGACCAACCT TGGGCAACAT AGCAAGACCC 2940  
CATTCTCTAA AAAAAAAAAA AAAAAATTTA AACTTAGCTG GGTATGGTGG CACATGCCTA 3000  
TAGTCTCAGC TACTTGTGAG GCTGAGGCAG GAGGATTCTT TGAGCCCAGG AGTTTGAGGT 3060  
5 TACAGTGAGC CACAATCACA CCAATCACTG CACTCCAGCC TGGGCAATAA AGTAACTCTT 3120  
GACTCAAAAA AATAAAAAAA ATTGTAGTGG TAGCCATGTG TTAATTGTTA AATAAATCTT 3180  
CCAAAGGGCT AAAAGTAAAT TACTTATAAA TTTTATATAG TTGTATTTT GACCTGCCTT 3240  
TTATATGTAT GAATATTTCA TAGTTTTGCA TATCAGATGT AGGCATACAG ACAAATACAT 3300  
10 AAACCAATGA ATATATTACA TATTCTGTGT TCCAATAAAA CTTTATTTAT GGACACTAAA 3360  
ATTTGAATTT CATAAAATTT TCCCATGTCA AGAATACAAA ATACTTGAGT TTTGTTTTTA 3420  
GCTATTTAAT AATAGGTCTC ATTTATTCCA CAGGCTGTAG TTTGTAGTCT TGCTTGAAAC 3480  
AATAGAAACA GACTGATTAA GCAGGAGAAG TTTTGTGAAA GAATTTTGTG TGGCTCACCG 3540  
AATTATTAGA AGGCAGGTGA ACCAGGAGGG TAAGCTTCCA GCAGCAATTT GTAAAACCAT 3600  
GCCTTAGAAT TGGACTAAGG AAGAAGCTGC TGACACTCCA CTGCCACACA GGGCACTGGA 3660  
15 AGAAAGTGTCT GCTGCCTCCC TGCCCCACCT TTGCCACTTC TGACAGCAGGA ATAGGTAGAA 3720  
GAATGCCCCC ACCCGCACCG GAACAGCAAC AAAAGGATTC TGCATGAGAT GCCTCCCTAA 3780  
ATTGCTGAAT TCAAAAAAGA AGTTGCATAC AAAGACATCT GATTGAAAAA GGGTATGTTA 3840  
TATGCCCTT TCATAGGCTG CTAGGGAGTT TTCTTGGTTC TACTTTCAGG TGGTGGGATC 3900  
AATAAGACCA GAATTTCTCA TATGTTGTGA GAGGATTCAA ATGTTACAGG GTTGCCAGCC 3960  
20 AAACATCAA TCATGTATAA ATCCAACAAA CACTTTGTAA CATAACAAGAA CTCAGGAAAT 4020  
GTGAACCAAT GTTGAGAAAT CTACTAAAAT CAGGCTTCCC GCAAACGAAG ATGAATGGAA 4080  
AATGTAAATA AAAAGAACTG GCAGTGTATA TCAGATGTTT AACTATAGGA CCAGAACTAA 4140  
GATGTGGAGA CTATTGCCAT AGACCACAAT GTAAATTTT AAGTGAGGAA GGAAAAATCA 4200  
GGAATCAAAA GGGGCCAGGT GCAGTGGCTC ACATCTATA TCCCAGAGCT TTGGGAGTTC 4260  
25 GAGGCAGGAG GATCACTTGA AGCCAGTTTT GAGACCAGCC TATGCAACAC ATTGAGACCC 4320  
TATCTCTACA AAAAATAGAT TAGCTGGGCA CCGTGGTGCA TGCCTATTGT CCTACCTACT 4380  
GTGGAGGCTG AAGTAGGAAA TCACTTGAGC CCGAGAGTTT GAGGTTACAG TGAGCTATGA 4440  
TTATACCACT GCACTCCAGC CTGGGCAAGA GAGCAAGACC TTGTCTCTT

Seq ID NO: 44 Protein sequence:  
Protein Accession #: CAB55276.2

1 11 21 31 41 51  
| | | | |  
35 MNGEYRGRGF GRGRFQSWKR GRGGGNFSGK WREREHRPDL SKTTGKRTSE QTPQFLLSTK 60  
TPQSMQSTLD RFIPYKGWKL YFSEVYSDSS PLIEKIQAFE KFFTRHIDLY DKDEIERKGS 120  
ILVDFKELTE GGEVTLNLPD IATELRDAPE KTLACMGLAI HQVLTKDLER HAAELQAQEG 180  
LSNDGETMVN VPHIHARVYN YEPLTQLKNV RANYYGKYIA LRGTVVRVSN IKPLCTKMAF 240  
40 LCAACGEIQS FPLPDGKYSL PTKCPVPVCR GRSFTALRSS PLTVTMDWQS IKIQELMSDD 300  
QREAGRIPRT IECELVDHLV DSCVPGDTVT ITGIVKVSNA BEGSRNKNDK CMFLLYIEAN 360  
SISNSKGQKT KSSEDGCKHG MLMEFSLKDL YAIQEIQAEE NLFKLIVNSL CPVIFGHELV 420  
KAGLALALFG GSQKYADDKN RPIRIGDPHI LVVGDPLGLK SQMLQAACNV APRGVYVCGN 480  
TTTTSGTLVT LSKDSSSGDF ALEAGALVLG DQGICGIDEF DKMGNHQHAL LEAMEQQSIS 540  
45 LAKAGVVCSL PARTSIIAAA NPVGGHYNKA KTVSENLMKG SALLSRFDLV FILDTPNEH 600  
HDHLLSEHVI AIRAGKQRTI SSATVARMNS QDSNTSVLEV VSEKPLSERL KVVPGETIDP 660  
IPHQLLRKYI GYARQYVYPR LSTEAAARVLS DFYLELRKQS QRLNSSPITT RQLESILRLT 720  
EARARLELRE EATKEDAEDI VEIMKYSMLG TYSDEFNLD FERSQHGSGM SNRSTAKRFI 780  
SALNNVAERT YNNIFQFHQL RQIAKELNIQ VADFENFIGS LNDQGYLLKK GPKVYQLQTM

Seq ID NO: 45 DNA sequence  
Nucleic Acid Accession #: NM\_005416.1  
Coding sequence: 149..658

1 11 21 31 41 51  
| | | | |  
55 ACCAGATCCC AGAGGCTGAA CACCTCGACC TTCTCTGCAC AGCAGATGAT CCCTGAGCAG 60  
CTGAAGACCA GAAAAGCCAC TAAGACTTTC TGCTTAATTC AGGAGCTTAG AGGATTCTTC 120  
60 AAAGAGTGTG TCCACGATCC TTTGAAGTAC GAGTTCTTAC CAGCAGAAGC AGACCTTTAC 180  
CCCACCACCT CAGCTTCAAC AGCAGCAGGT GAAACAACCC AGCCAGCCTC CACCTCAGGA 240  
AATATTTGTT CCCACAACCA AGGAGCCATG CCACTCAAAG GTTCCACAAC CTGGAAACAC 300  
AAAGATTCCA GAGCCAGGCT GTACCAAGGT CCCTGAGCCA GGCTGTACCA AGGTCCCTGA 360  
GCCAGGCTGT ACCAAGGTCC CTGAGCCAGG TTGTACCAAG GTCCCTGAGC CAGGCTGTAC 420  
CAAGGTCCCT GAGCCAGGTT GTACCAAGGT CCCTGAGCCA GGCTACACCA AGGTCCCTGA 480  
65 ACCAGGCAGC ATCAAGGTCC CTGACCAAGG CTTCTCAAG TTTCTGAGC CAGGTGCCAT 540  
CAAAGTTCCT GAGCAAGGAT ACACCAAGT TCCTGTGCCA GGCTACACAA AGCTACCAGA 600  
GCCATGTCCT TCAACGGTCA CTCCAGGCCC AGCTCAGCAG AAGACCAAGC AGAAGTAATT 660  
TGGTGCACAG ACAAGCCCTT GAGAAGCCAA CCACCAGATG CTGGACACCC TCTTCCCATC 720  
70 TGTTTCTGTG TCTTAATTGT CTGTAGACCT TGTAATCAGC ACATTGTAC CCAAGCCAT 780  
AGTCTCTCTC TTATTTGTAT CCTAAAAATA CGTACTATAA AGCTTTTGT CACACACACT 840  
CTGAAGAATC CTGTAAGCCC CTGAATTAAG CAGAAAGTCT TCATGGCTTT TCTGGTCTTC 900  
GGCTGCTCAG GGTTCATCTG AAGATTCGAA TGAAAAGAAA TGCATGTTTC CTGCTCTTCC 960  
CTCATTAAT TGCTTTTAAT TCCA

Seq ID NO: 46 Protein sequence:  
Protein Accession #: NP\_005407.1

1 11 21 31 41 51  
| | | | |  
80 MSSYQQKQTF TPPPQLQQQQ VKQPSQPPPQ EIFVPTTKEP CHSKVPQPGN TKIPEPGCTK 60  
VPEPGCTKVP EPGCTKVPEP GCTKVPEPGC TKVPEPGCTK VPEPGYTKVP EPGSIKVPDQ 120  
GFIKFPEPGA IKVPEQGYTK VPVPGYTKLP EPCPSTVTPG PAQQKTKQK

Seq ID NO: 47 DNA sequence  
Nucleic Acid Accession #: Eos sequence



1 11 21 31 41 51  
| | | | |  
5 GCGTCGTGTG CAGGCGTCCC CGGGCTGTGG ATAATTAGAC ACGTTCTTCC CTCATTGCCC 60  
AAGGCTCGTT AGAATTCGCC CTAGAGCTGT ATCATGTATT TTCTTTCAA TTAACCTTGC 120  
TTGCAATTAA GCTTAGGGAA CCAGCAACAA AAGCAAACCTT GGCCCGAGGT CGTTCACCGC 180  
GAAAATGGAT TAGAGAAACT TCTTCCCCGA TTTAAGGGGA AAGATTCCTG CGGCCAGCGC 240  
TTTGGGGAAG GTGCCCCGAC CGCAGAGGCG ACGACAGGGG AGCAGGAAGC TGCTCACGGT 300  
AGTCGGCGTT GGCGGCAGCG GTGGCCTTCC TCATCTGGGC GATGTGGGCT CCTAGAAGAG 360  
10 TAAGGATAAC ATCCTGGAAA TGACTTCTGT ACGGTTTGAG CCCAACTGCA CACTCATGAC 420  
TTGGAGCTGC CCTGTGGAGT TACAGTTTAC CAAACACATT CATGAACATA ATCTCATTTA 480  
CTAAAAACTT TGTGAGAATT TTCTTTTACT AAAATTTTTT CTTATTACAA A

Seq ID NO: 48 DNA sequence:  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
| | | | |  
20 TTCCAAATTT TTTTITTTGT AATAAGAAAA AATTTTAGTA AAAGAAAATT CTCACAAAGT 60  
TTTTAGTAAA TGAGATTATG TTCATGAATG TGTTTGGTAA ACTGTAACCT CACAGGGCAG 120  
CTCCAAGTCA TGAGTGTGCA AACCCTACAG AAGTCATTTC CAGGATGTTA 180  
TCCTTACTCT TCTCGGAGCC CACATCGCCC AGATGAGGAA GGCCACCGCT GCCGCCAACG 240  
CCGACTACCG TGAGCAGCTT CCTGCTCCCC TGTCGTGCGC TCTGCGGTCG GGGCACTTTC 300  
25 CCCAAAGCGC TGGCCGAGG AATCTTTCCC CTTAAATCGG GGAAGAAGTT TCTCTAATCC 360  
ATTTTCGCGG TGAACGACCT CGGGCCAAGT TTGCTTTTGT TGCTGGTTCC CTAAGCTTAA 420  
TTGCAAGCAA AGTTAATTTG AAAGAAAATA CATGATACAG CTCTAGGGCG AATTCTAACG 480  
AGCCTTGGGC AATGAGGGAA GAACGTGTCT AGTTATCCAC AGCCCGGGGA CGCCTGCACA 540  
CGACGCT

Seq ID NO: 49 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
| | | | |  
35 TCTTTCTTCT GCTGCTCGTT TGTCTCTCCT GTGCTCTTCT TCTTTCTTTC CCTCGCCGCT 60  
CCTGCCGACC TCTGTTGTCT CTTCTCTGAT GGCGGGGGGC GGGAGAAGCT GACCGGTGAG 120  
ACCGTAGACC CGAAACCATT GGGTGTCAAC AGCCGGTTCGC CGGCTTTTTT GGGAGAACCC 180  
GACACATGCA GACCAGTTTT CCTGGAACNG CATGACCATG TTATTACTAT GGGCCGCCTC 240  
40 CCCAACCAAA GTGTTTAAAA CTTTTTAGGG CACCCCAAAA ATTTTTTTTT TTTTTTTTTT 300  
TTCAATTAAA AAAGTCTAAT ATTTATATTA AATACAAAGA TACCCAAACC CTTTATGCTT 360  
CTTTCTCTGA TCTGTGCTT TTTTCTTTGA CAGCATCTCC ATTTTTTTTC TGCTGCTTCA 420  
TCGCTGTAGC CATGGGAATC CGTTTCATTA TTATGGTAGC AATATGGAGT GCTGTATTCC 480  
TAAAGAAACT GACACAGGAG AATCACTTGA ACTTGGGAGG CAGAGTTTGC AGTGAGCCGA 540  
GATTGAACCA GTGCACTCCA GCCTTGGCAG CGGAGCAAGA TTCTGTCAAC GTTCCTGAAG 600  
45 TGCTGGTATC GTCCTGCAGC CCCATCCTCG GTTCCATTGC GCTGCCAGGC AGGGTGCTGG 660  
GACGTGGGGA GAGCTGGTCT ATATATCCGG GTGAAGCTCA GCTGTGGCAC ACCTTGATG 720  
CCGGGTCTCT CTTGGCCCCG GGGACCTAGT ATTTTGTCCA CGAGTGTACA CCAAACAAAG 780  
GAGACAGCAT CATTTATGAG CCTGCAGCAT CCACCCTACT GCTGTATCCA GTTTCATTG 840  
ACTG

Seq ID NO: 50 DNA sequence  
Nucleic Acid Accession #: L05187  
Coding sequence: 1991..2260

1 11 21 31 41 51  
| | | | |  
55 CTGCAGGGAG GCAGGTAGAA AAGGCTTTTG GGTTTTTCAGG TGGGGGGCAG TCTAGCCTGA 60  
TCAGAAAGGA GGAAGAGGCC AGGGCAGATG TCTGGGTGGA GTGAAGGGAA AAAGTGATCC 120  
60 CAGAAGAAGG ATTAGCCCCC GAAAGTCCCT GAAAGTAGAG AAGGGTAAAG GTGTGGTTGG 180  
TGAAGGAAAG CAGGTTTTTC CAGATTAGCA ACCAGTCAGG GGGAGGAAGG TGAGAGTGGG 240  
AGAGTCATAA GTAAATTATT CTGAATGTGT GTAGTTTAAAT GGAATTGGGA AAAAGATGGG 300  
GGAAATGGAT GGAAGGTCTT GGAATCTGAG ACAAGGGGTC TATAATCAGT CCATTTTCATT 360  
ATTTCTAGCT TCCACCTTCA CCAAGGCAGA CAAGGAGGGC CCACCTCAGC TCCTCTGCTC 420  
65 CCCCTCCCTT TCCACCTAT TCATGTGTGC AAGAGTGGCC TGTCCACAG AACACGGGGA 480  
ACAACCATCT CAATGACAAG GACAGCAGGT GGCAAGGCTC AACAGGACTC AGATGTCCCC 540  
CCAGGGTTAA CTCATGAAAC CCTCCATGAA GCCTGCTGCT CACCCCTCCC TCAAGGCAAG 600  
CCCTGCACCT GGGTCTGAGG ATGAGGGTGG CAGTGAAGAT TAGGCCAGTG ACATCATTTT 660  
CAGCCAGCTA GTGCCAAAA ATATCAGGTG GTGTTCATCA AATAAGCCGA GCCAACCAGT 720  
70 GATGAGGATG GTAGTGTGAG TCATGTGTGA CAGGTGAGGA ATGAAAAACAG AGTGCCCGAG 780  
AGCTTCTATT TCCTTGAGGC AGGGCTCATT CATCTTATAA AAGCCAGCTG GCCATTGCCT 840  
TCACACCAAA CCCAAGGGAC CACACAGCCC ATTTCTGCTC GTATACCAG TAAGTCTCTG 900  
ATTGCAACAA ACTGGCAATT CTAGTGTACT TTTTCATTAT TAGAAATTAG CTAAAGGCAA 960  
ATATGTGTAA GCAGGTTAAT CCAGGGTTTC AATGGGAGAT AGAGAATAGT GGAATATCTT 1020  
75 TATTTTAAGT TAAATTACAG TCTGGATTGG AAAGGACCTT AGAGATGGTT AGGGCTCCCA 1080  
CCTCAGTAGA TAGTCATTGA ACTGGGAGTC CTGGAGAAGA TTGTTCAAAT GCCCATGGGA 1140  
AGTTCATAGC AGAACTAGAA CTCAGGCCAG AGCACTCTCA GTAACACTGC AATTTCCCCC 1200  
TGACAAGATA TTTATAGAAA TTTTAATTTA TTAGATGGAT CTCTACTGAG CATTTATTCC 1260  
ATTTAAGGCA GTATGCTAGG CACTTTGGAC AAATCAATGC CCTAACGTAC TTACTTAACA 1320  
AACATAAAAC CTAGCAGGAA GGTAATACAT ATATATAAAT AAATGAAATG CAAAGTAGAT 1380  
80 AGTAATTGGC ATGACGGAGA TGGGCAGAGA AGGGCTGTGC ACTTTTGGGA GACTTGCTCA 1440  
AGGAGACCTC TAGGGTGTCA AGTGATGTGA GCTATGATGG AGGGGTATTT GGACAAGCAG 1500  
AGATGGGAAG AAAAGCATTT GGAAGGGACT GTGTAAGCAC AGACCAGAAG CAAAACCATA 1560  
GAGGCTTAGA TGAATATAAA GCCATCCTAT AAGTCACAGG CTTTCTACAT GGTACTAGGA 1620  
GAGGAAAGTG GTTCTGATGCC ATTTTCCAAA AGACCTAATA TGCGGACCTC ATGTCCCTCA 1680  
85 GAAGCCAGCT TTAGTAGGGC ATTTTTCAG AACAGATATA AGGTGCCTTG GGTAGGAAGG 1740  
GAGCCAAGAA GAGAACTCCA ATAAATGGA GCAGAAGAAA TTGCCTTTTA GCTCCTCCTC 1800  
TTCAAAGGGC CTGAAAATTA TCCAAGCTTA TTTTATTTTT AAATGTAATG GGGGAGCTAA 1860



5 GGGAGATGAA AGGCTTTCTC TTCTAAAGGG TCCTGAAATA AAATCTGTTT GGCATTGAAT 1920  
TTGTATCCAT CTTTCTTTAA TTGAATCACT GTGTCAGCTT TCTGTCTCTA GAAAAAACA 1980  
CATTGGAAGC ATGAATTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA 2040  
GCAGCAGCAG GTGAAACAAC CTTGCCAGCC TCCACCCAG GAACCATGCA TCCCCAAAAC 2100  
CAAGGAGCCC TGCCAACCCA AGGTGCCTGA GCCCTGCCAC CCCAAAGTGC CTGAGCCCTG 2160  
CCAGCCCAAG ATTCCAGAGC CCTGCCAGCC CAAGGTGCCT GAGCCCTGCC CTTCAACGGT 2220  
CACTCCAGCA CCAGCCCAGC AGAAGACCAA GCAGAAGTAA TGTGGTCCAC AGCCATGCCC 2280  
TTGAGGAGCT GGCCACTGGA TACTGAACAC CCTACTCCAT TCTGCTTATG AATCCCATT 2340  
10 GCCTATTGAC CCTGCAGTTA GCATGCTGTC ACCCTGAATC ATAATCGCTC CTTTGCACCT 2400  
CTAAAAAGAT GTCCCTTACC CTCATTCTGG AGGCTCCTGA GCCTCTGCGT AAGGCTGAAC 2460  
GTCTCACTGA CTGAGCTAGT CTTCTTGTG CTCGGGTGCA TTTGAGGATG GATTTGGGGA 2520  
AGGTCAAGTG ACCATCCCTA G

15 Seq ID NO: 51 Protein sequence:  
Protein Accession #: AAC26838

1 11 21 31 41 51  
| | | | | |  
20 MNSQQQKQPC TPPPQPPQQQ VKQPCQPPPO EPCIPKTKEP CQPKVPEPCH PKVPEPCQPK 60  
IPEPCQPKVP EPCPSTVTPA PAQQKTKQK

25 Seq ID NO: 52 DNA sequence  
Nucleic Acid Accession #: NM\_002638.1  
Coding sequence: 120-473

1 11 21 31 41 51  
| | | | | |  
30 CAATACAGCT AAGGAATTAT CCCTTGTAAT TACCACAGAC CCGCCCTGGA GCCAGGCCAA 60  
GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120  
TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG GTTTCCTCAT CGCTGGGACG CTGGTTCTAG 180  
AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCCTAT 240  
TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300  
35 CGCAAGAGCC AGTCAAAGGT CCAGTCTCCA CTAAGCCTGG CTCCTGCCCC ATTATCTTGA 360  
TCCGGTGCGC CATGTTGAAT CCCCCTAACC GCTGCTTGA AGATACTGAC TGCCAGGAA 420  
TCAAGAAGTG CTGTGAAGGC TCTTGCGGGA TGGCCTGTT CGTTCCCCAG TGAAGGGAGC 480  
CGGTCTCTTG TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540  
TGCTGCCCTT CCCCTTCCCA CACTGTCCAT TCTTCCTCCC ATTCAGGATG CCCACGGCTG 600  
40 GAGCTGCCTC TCTCATCCAC TTTCCAATAA A

Seq ID NO: 53 Protein sequence:  
Protein Accession #: NP\_002629.1

1 11 21 31 41 51  
| | | | | |  
45 MRASSFLIVV VFLIAGTLVL BAAVTGVVPK GQDTVKGKRV FNGQDPVKQG VSVKGQDKVK 60  
AQEPVKGPVS TKPGSCPIIL IRCAMLNPPN RCLKDTPCPG IKKCCGSCG MACFVPQ

50 Seq ID NO: 54 DNA sequence  
Nucleic Acid Accession #: NM\_019618  
Coding sequence: 75-584

1 11 21 31 41 51  
| | | | | |  
55 GGCACGAGCC ACGATTCACT CCCCTGGACT GTAGATAAAG ACCCTTTCTT GCCAGGTGCT 60  
GAGACAACCA CACTATGAGA GGCACCTCCAG GAGACGCTGA TGGTGGAGGA AGGGCCGTCT 120  
ATCAATCAAT GTGTAAACCT ATTACTGGGA CTATTAATGA TTTGAATCAG CAAGTGTGGA 180  
CCCTTCAGGG TCAGAACCTT GTGGCAGTTC CACGAAGTGA CAGTGTGACC CCAGTCACTG 240  
60 TTGCTGTTAT CACATGCAAG TATCCAGCA CTCTTGAGCA AGGCAGAGGG GATCCCATT 300  
ATTTGGGAAT CCAGAATCCA GAAATGTGTT TGTATTGTGA GAAGGTTGGA GAACAGCCCA 360  
CATTCGAGCT AAAAGAGCAG AAGATCATGG ATCTGTATGG CCAACCCGAG CCCGTGAAAC 420  
CCTTCCTTTT CTACCGTGCC AAGACTGGTA GGACCTCCAC CCTTGAGTCT GTGGCCTTCC 480  
CGGACTGGTT CATTCGCTCC TCCAAGAGAG ACCAGCCCAT CATTCTGACT TCAGAACTTG 540  
GGAAGTCATA CAACACTGCC TTTGAATTAA ATATAAATGA CTGAACTCAG CCTAGAGGTG 600  
65 GCAGCTGGT CTTTGTCTTA AAGTTTCTGG TTCCCAATGT GTTTTCGTCT ACATTTTCTT 660  
AGTGTCAATT TCACGCTGGT GCTGAGACAG GGGCAAGGCT GCTGTTATCA TCTCATTTTA 720  
TAATGAAGAA GAAGCAATTA CTTCATAGCA ACTGAAGAAC AGGATGTGGC CTCAGAAGCA 780  
GGAGAGCTGG GTGGTATAAG GCTGTCTCT CAAGCTGGTG CTGTGTAGGC CACAAGGCAT 840  
70 CTGCATGAGT GACTTTAAGA CTCAAAGACC AAACACTGAG CTTTCTTCTA GGGGTGGGTA 900  
TGAAGATGCT TCAGAGCTCA TCGCGTTTAC CCACGATGGC ATGACTAGCA CAGAGCTGAT 960  
CTCTGTTTCT GTTTTGTCTT ATTCCCTCTT GGGATGATAT CATCCAGTCT TTATATGTTG 1020  
CCAATATACC TCATTGTGTG TAATAGAACC TTCTTAGCAT TAAGACCTTG TAAACAAAAA 1080  
TAATCTTGT GTTAAGTTAA ATCATTTTGT TCCTAATTGT AATGTGTAAT CTTAAAGTTA 1140  
75 AATAAACTTT GTGTATTTAT ATAATAAAAA AAAAAAAAAA AAA

Seq ID NO: 55 Protein sequence:  
Protein Accession #: NP\_062564

1 11 21 31 41 51  
| | | | | |  
80 MRGTPGDADG GGRAVYQSMC KPITGTINDL NQQVWTLQOQ NLVAVPRSDS VTPVTVAVIT 60  
CKYPEALEQG RGDPIYLGIIQ NPEMCLYCEK VGEQPTLQLK EQKIMDLYGQ PEPVKPFLFY 120  
RAKTGRTSTL ESVAFPDWFI ASSKRDQPII LTSELGKSYN TAFELNIND

85 Seq ID NO: 56 DNA sequence  
Nucleic Acid Accession #: NM\_003125  
Coding sequence: 65-334



1 11 21 31 41 51  
| | | | |  
5 AGCAGTTCTA AGGGACCATA CAGAGTATTC CTCTCTTCAC ACCAGGACCA GCCACTGTTG 60  
CAGCATGAGT TCCCAGCAGC AGAAGCAGCC CTGCATCCCA CCCCCTCAGC TTCAGCAGCA 120  
GCAGGTGAAA CAGCCTTGCC AGCCTCCACC TCAGGAACCA TGCATCCCA AAACCAAGGA 180  
GCCCTGCCAC CCAAGGTGC CTGAGCCCTG CCACCCCAA GTGCCCTGAGC CCTGCCAGCC 240  
CAAGCTTCCA GAGCCATGCC ACCCCAAGGT GCCTGAGCCC TGCCCTTCAA TAGTCACTCC 300  
AGCACCAGCC CAGCAGAAGA CCAAGCAGAA GTAATGTGGT CCACAGCCAT GCCCTTGAGG 360  
10 AGCCGGCCAC CAGATGCTGA ATCCCTATC CCATTCTGTG TATGAGTCCC ATTTGCTTG 420  
CAATTAGCAT TCTGTCTCCC CCAAAAAAGA ATGTGCTATG AAGCTTTCTT TCCTACACAC 480  
TCTGAGTCTC TGAATGAAGC TGAAGGTCTT AGTACCAGAG CTAGTTTCA GCTGCTCAGA 540  
ATTCATCTGA AGAGAGACTT AAGATGAAAG CAAATGATTC AGCTCCCTTA TACCCCATTT 600  
15 AAATTCACCTT TCAATTCCA

Seq ID NO: 57 Protein sequence:  
Protein Accession #: NP\_003116

1 11 21 31 41 51  
| | | | |  
20 MSSQQQKQPC IPPPQLQQQQ VKQPCQPPPO EPCIPKTKEP CHPKVPEPCH PKVPEPCQPK 60  
25 LPEPCHPKVP EPCPSIVTPA PAQQKTKQK

Seq ID NO: 58 DNA sequence  
Nucleic Acid Accession #: NM\_001793.2  
Coding sequence: 71-2560

1 11 21 31 41 51  
| | | | |  
30 AAAGGGGCAA GAGCTGAGCG GAACACCGGC CCGCCGTGCG GGCAGCTGCT TCACCCCTCT 60  
CTCTGCAGCC ATGGGGCTCC CTCGTGGACC TCTCGCGTCT CTCCTCCTTC TCCAGGTTTG 120  
35 CTGGCTGCAG TGCGCGGCCT CCGAGCCGTG CCGGCGGTC TTCAGGGAGG CTGAAGTGAC 180  
CTTGGAGGCG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAAG TATTCATGGG 240  
CTGCCCTGGG CAAGAGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300  
TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360  
ATCCAAACGT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420  
40 TGAAAAATGGC AAGGGTCCCT TCCCCCAGAG ACTGAATCAG CTCAAGTCTA ATAAAGATAG 480  
AGACACCAAG ATTTTCTACA GCATCACGGG GCCGGGGGCA GACAGCCCCC CTGAGGGTGT 540  
CTTCGCTGTA GAGAAGGAGA GAGGCTGGTT GTTGTGAAT AAGCCACTGG ACCGGGAGGA 600  
GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTGAGAG AATGGTGCCT CAGTGGAGGA 660  
CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720  
45 GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCTACCA GGTACTTCTG TGATGCAGGT 780  
GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAT GGGGTGGTTG CTTACTCCAT 840  
CCATAGCCAA GAACCAAAGG ACCCACACGA CCTCATGTTT ACCATTACC GGAGCACAGG 900  
CACCATCAGC GTCATCTCCA GTGGCCTGGA CCGGGAAAAA GTCCCTGAGT ACACACTGAC 960  
CATCCAGGCC ACAGACATGG ATGGGGACGG CTCCACCACC ACGGCAGTGG CAGTAGTGGA 1020  
GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080  
50 GCCTGAGAAT GCAGTGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140  
CAACTACCA GCGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACG GGGACCATTT 1200  
TACCATCACC ACCCACCCTG AGAGCAACCA GGGCATCCTG ACAACCAGGA AGGGTTTGA 1260  
TTTTGAGGCC AAAAACCAGC ACACCTGTA CGTTGAAGTG ACCAACGAGG CCCCTTTTGT 1320  
55 GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAC GTGGAGGATG TGAATGAGGC 1380  
ACCTGTGTTT GTCCCACCCT CCAAAGTCGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440  
GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCAGCTACCG 1500  
CATCCTGAGA GACCCAGCAG GGTGGCTAGC CATGGACCCA GACAGTGGGC AGGTACACAGC 1560  
TGTGGGCACC CTCGACCGTG AGGATGAGCA GTTTGTGAGG AACAACATCT ATGAAGTCAT 1620  
60 GGTCTTGGCC ATGGACAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGCTAAC 1680  
ACTGATTGAT GTCAATGACC ATGGCCAGT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740  
CCAAAGCCCT GTGCGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800  
CCCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860  
GGAAGGTGAC ACAGTGGTCT TGTCCCTGAA GAAGTTCTCT AAGCAAGATA CATATGACGT 1920  
GCACCTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980  
65 GTGCGACTGC CATGGCCATG TCGAAACCTG CCCTGGACCC TGGAAGGGAG GTTTTCTCCT 2040  
CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT TGTCTCTCTG CTGGTGCTGC TTTTGTGTTG 2100  
GAGAAAGAAG CGGAAGATCA AGGAGCCCCT CTTACTCCCA GAAGATGACA CCCGTGACAA 2160  
CGTCTTCTAC TATGGCGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220  
70 GCTCCACCGA GGTCTGGAGG CCAGGCCGGA GGTGTTCTC CGCAATGACG TGGCACCAAC 2280  
CATCATCCCG ACACCCATGT ACCGTCTCTG GCCAGCCAAC CCAGATGAAA TCGGCAACTT 2340  
TATAATTGAG AACCTGAAGG CGGCTAACAC AGACCCACA GCCCCGCCCT ACGACACCTT 2400  
CTTGGTGTTC GACTATGAGG GCAGCGGCTC CGACGCGCG TCCCTGAGCT CCCTCACCTC 2460  
CTCCGCCCTC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCCGCTTCAA 2520  
75 GAAGCTGGCA GACATGTACG GTGGCGGGGA GGACGACTAG GCGGCCTGCC TGCAGGGCTG 2580  
GGGACCAAAC GTCAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG 2640  
GACTTCGAG CTTGTGAGGA AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700  
ACGTTAGAGT GGTGCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTTGACTTC 2760  
AGCACTGAAA ACCTTCCAC CTGGGCCAGG GTTGCTCAG AGGCCAAGTT TCCAGAAGCC 2820  
80 TCTTACTGTC CGTAAATGTC TCAACCCTGT GTCTGGGCC TGGGCTGCT GTGACTGACC 2880  
TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTTT 2940  
TTTTTTTAAAT GCTATCTTCA AAACGTTAGA GAAAGTTCTT CAAAAGTGCA GCCCAGAGCT 3000  
GCTGGGCCCA CTGGCCGTCC TGCATTTCTG GTTTCCAGAC CCCAATGCCT CCCATTCGGA 3060  
TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGGTTGCC CCTTATTTTT TATTTTCCCT 3120  
85 GTTGCCTTGC TATAGATGAA GGGTGAAGG AATCGTGTAT ATGTACTAGA ACTTTTTTAT 3180  
TAAAGAACT TTTCCAGAA AAAAA

Seq ID NO: 59 Protein sequence:



Protein Accession #: NP\_001784.2

	1	11	21	31	41	51	
5							
	MGLPRGPLAS	LLLLQVCWLQ	CAASEPCRAV	FREAEVTLEA	GGAEQEPGQA	LGKVFMGCPG	60
	QEPALFSTDN	DDFTVRNGET	VQERRSLKER	NPLKIFPSKR	ILRRHKRDWV	VAPISVPENG	120
	KGPFQRLNQ	LKSNKDRDTK	IFYSITGPGA	DSPPEGVFAV	EKETGWLLLN	KPLDREEIAK	180
	YELFGHAVSE	NGASVEDPMN	ISIIIVTDQND	HKPKFTQDTF	RGSVLEGVLP	GTSVMQVTAT	240
10	DEDDAIYTYN	GVVAYSIHSQ	EPKDPHDLMF	TIHRSTGTIS	VISSGLDREK	VPEYTLTIQA	300
	TDMDGDGSTT	TAVAVVEILD	ANDNAPMFDP	QKYEAHVPEN	AVGHEVQRLT	VTDLDPNSP	360
	AWRATYLMG	GDDGDHFTIT	THPESNQGIL	TTRKGLDFEA	KNQHTLYVEV	TNEAPFVLKL	420
	PTSTATIVVH	VEDVNEAPVF	VPPSKVVEVQ	EGIPTGEPVC	VYTAEDPDKE	NQKISYRILR	480
	DPAGWLAMP	DSGQVTAVGT	LDREDEQFVR	NNIYEVMLA	MDNGSPPTTG	TGTLTLLTID	540
	VNDHGPVPEP	RQITICNQSP	VRQVLNITDK	DLSPHTSFQ	AQLTDDSDIY	WTAEVNEEGD	600
15	TVVLSLKKFL	KQDTYDVHLS	LSDHGNKEQL	TVIRATVCD	HGHVETCPGP	WKGGFILPVL	660
	GAVLALLFL	LVLLLLLVRKK	RKIKEPLLLP	EDDTRDNVY	YGEEGGGEED	QDYDITQLHR	720
	GLEARPEVVL	RNDVAPTIIP	TPMYRPRPAN	PDEIGNFIE	NLKAANTDPT	APPYDTLLVF	780
	DYEGSGSDAA	SLSSLTSSAS	DQDQDYDYL	EWGSRFKKLA	DMYGGGEDD		

20 Seq ID NO: 60 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 162-428

	1	11	21	31	41	51	
25							
	GCGTTCGGTT	GGCGGCGGAT	TGCAACGTTT	GGACTGAGGT	TTTTCTGCCT	GAAGAAGCGT	60
	CATACGGACC	GGATTGTTTT	CGCTGGCCCA	TGCTCCCGG	AGCTTGCTGT	CGATACAGAG	120
	AGCACCTCGG	AAGCTGAGGC	AGCTGGTACT	TGACAGAGAG	GATGGCGCTG	TGACCATAG	180
30	TCTCCAGAG	GAAGCAGATA	AAGCGGAAGG	CTCCCCGTGG	CTTTCTAAAG	CGAGTCTTCA	240
	AGCGAAAGAA	GCCTCAACTT	CGTCTGGAGA	AAAGTGGTGA	CTTATTGGTC	CATCTGAAC	300
	GTTTACTGTT	TGTTTCATCGA	TAGCAGAAAG	AGTCCAGGAC	AAACGCTTGT	GCGAGTAAAT	360
	GTAGAGTCAT	TAACAAGGAG	CATGTACTGG	CCGCAGCAAA	GGTAATTCTA	AAGAAGAGCA	420
	GAGGTTAGAA	GTCAAAGAAC	ATATTCTTGA	AAGTTATGAT	GCATTCTTTT	GGGTGGTAAC	480
35	AGATCATAAA	GACATTTTTT	ACACATCAGT	TAATATGGGA	TTATTAAATA	TTGG	

Seq ID NO: 61 Protein sequence:  
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
40							
	MALSTIVSQR	KQIKRKAPRG	FLKRVFKRKK	PQLRLEKSGD	LLVHLNCLLF	VHRLAESRT	60
	NACASKCRVI	NKEHVLAAAK	VILKKSRG				

45 Seq ID NO: 62 DNA sequence  
Nucleic Acid Accession #: NM\_000094.2  
Coding sequence: 99-8933

	1	11	21	31	41	51	
50							
	GGGCTGGAGG	GGCGCTGGGC	TGGGACCTGC	CAAGGCCACC	GCAGGGGGGA	GCAAGGGACA	60
	GAGGCGGGGG	TCCTAGCTGA	CGGCTTTTAC	TGCCTAGGAT	GACGCTGCGG	CTTCTGGTGG	120
	CCGCGCTCTG	CGCCGGGATC	CTGGCAGAGG	CGCCCCGAGT	GCGAGCCCAG	CACAGGGAGA	180
	GAGTGACCTG	CACGCGCCTT	TACGCGCTG	ACATTGTGTT	CTTACTGGAT	GGCTCCTCAT	240
55	CCATTGGCCG	CAGCAATTTT	CGCGAGGTCC	GCAGCTTTCT	CGAAGGGCTG	GTGCTGCCTT	300
	TCTCTGGAGC	AGCCAGTGCA	CAGGGTGTGC	GCTTTGCCAC	AGTGCAGTAC	AGCGATGACC	360
	CACGGACAGA	GTTCCGGCCT	GATGCACTTG	GCTCTGGGGG	TGATGTGATC	CGCGCCATCC	420
	GTGAGCTTAG	CTACAAGGGG	GGCAACACTC	GCACAGGGGC	TGCAATTCTC	CATGTGGCTG	480
	ACCATGTCTT	CCTGCCCCAG	CTGGCCCCAG	CTGGTGTCCT	CAAGGTCTGC	ATCCTGATCA	540
60	CAGACGGGAA	GTCCAGGAC	CTGGTGACAC	GTGAGCGGCA	AAGGCTGAAG	GGGCAGGGGG	600
	TCAAGCTATT	TGCTGTGGGG	ATCAAGAATG	CTGACCCTGA	GGAGCTGAAG	CGAGTTGCCT	660
	CACAGCCAAC	CTCCGACTTC	TTCTTCTTCG	TCAATGACTT	CAGCATCTTG	AGGACACTAC	720
	TGCCCCCTCG	TTCCCGGAGA	GTGTGCACGA	CTGCTGGTGG	CGTGCTGTGT	ACCCGACCTC	780
	CGGATGACTC	GACCTCTGCT	CCACGAGACC	TGGTGTCTGT	TGAGCCAAGC	AGCCAATCCT	840
	TGAGAGTACA	GTGGACAGCG	GCCAGTGGCC	CTGTGACTGG	CTACAAGGTC	CAGTACACTC	900
65	CTCTGACGGG	GCTGGGACAG	CCACTGGCGA	GTGAGCGGCA	GGAGGTGAAC	GTCCCAGCTG	960
	GTGAGACCAG	TGTGCGGCTG	CGGGGTCTCC	GGCCACTGAC	CGAGTACCAA	GTGACTGTGA	1020
	TTGCCCTCTA	CGCCAACAGC	ATCGGGGAGG	CTGTGAGCGG	GACAGCTCGG	ACCACTGCCC	1080
	TAGAAGGGCC	GGAAGTACC	ATCCAGAATA	CCACAGCCCA	CAGCCTCCTG	GTGGCCTGGC	1140
70	GGAGTGTGCC	AGGTGCCACT	GGCTACCGTG	TGACATGGCG	GGTCTCAGT	GGTGGGCCCA	1200
	CACAGCAGCA	GGAGCTGGGC	CCTGGGCAGG	GTTCAGTGT	GCTGCGTGAC	TTGGAGCCTG	1260
	GCACGAGCTA	TGAGGTGACC	GTGAGTACCC	TATTTGGCCG	CAGTGTGGGG	CCCCCACTT	1320
	CCCTGATGGC	TGCACTGAC	GCTTCTGTTG	AGCAGACCCT	GCGCCCGGTC	ATCCTGGGCC	1380
	CCACATCCAT	CCTCTTTTCC	TGGAAGTTGG	TGCCTGAGGC	CCGTGGCTAC	CGGTTGGAAT	1440
75	GGCGGCGTGA	GACTGGCTTG	GAGCCACCGC	AGAAGGTGGT	ACTGCCCTCT	GATGTGACCC	1500
	GCTACCAATT	GGATGGGCTG	CAGCCGGGCA	CTGAGTACCG	CCTCACACTC	TACACTCTGC	1560
	TGGAGGGCCA	CGAGGTGGCC	ACCCCTGCAA	CCGTGGTTCC	CACTGGACCA	GAGCTGCCTG	1620
	TGAGCCCTGT	AACAGACCTG	CAAGCCACCG	AGCTGCCCGG	GCAGCGGGTG	CGAGTGTCTT	1680
	GGAGCCCACT	CCCTGGTGCC	ACCCAGTACC	GCATCATTTG	GCGCAGCACC	CAGGGGGTTG	1740
80	AGCGGACCCCT	GGTGCTTCCT	GGGAGTCAGA	CAGCATTCGA	CTTGGATGAC	GTTCAGGCTG	1800
	GGCTTAGCTA	CACTGTGCGG	GTGTCTGCTC	GAGTGGGTCC	CCGTGAGGGC	AGTGCCAGTG	1860
	TCCTCACTGT	CCGCGGGGAG	CCGGAAACTC	CACTTGCTGT	TCCAGGGCTG	CGGGTTGTGG	1920
	TGTCAGATGC	AACGCGAGTG	AGGGTGGCCT	GGGGACCCGT	CCCTGGAGCC	AGTGGATTTT	1980
	GGATTAGCTG	GAGCACAGGC	AGTGGTCCGG	AGTCCAGCCA	GACACTGCCC	CCAGACTCTA	2040
	CTGCCACAGA	CATCACAGGG	CTGCACCTTG	GAACCACTTA	CCAGGTGGCT	GTGTCCGTAC	2100
85	TGCGAGGCAG	AGAGGAGGGC	CCTGCTGCAG	TCATCGTGGC	TCGAACGGAC	CCACTGGGCC	2160
	CAGTGAGGAC	GGTCCATGTG	ACTCAGGCCA	GCAGCTCATC	TGTCACCAT	ACCTGGACCA	2220
	GGGTTCTCTG	CGCCACAGGA	TACAGGGTTT	CCTGGCACTC	AGCCACGGC	CCAGAGAAAT	2280



	CCCAGTTGGT	TTCTGGGGAG	GCCACGGTGG	CTGAGCTGGA	TGGACTGGAG	CCAGATACTG	2340
	AGTATACGGT	GCATGTGAGG	GCCCATGTGG	CTGGCGTGGA	TGGGCCCCCT	GCCTCTGTGG	2400
	TTGTGAGGAC	TGCCCCTGAG	CCTGTGGGTC	GTGTGTCGAG	GCTGCAGATC	CTCAATGCTT	2460
5	CCAGCGACGT	TCTACGGATC	ACCTGGGTAG	GGGTCACTGG	AGCCACAGCT	TACAGACTGG	2520
	CCTGGGGCCG	GAGTGAAGGC	GGCCCCATGA	GGCACCAGAT	ACTCCCAGGA	AACACAGACT	2580
	CTGCAGAGAT	CCGGGGTCTC	GAAGGTGGAG	TCAGCTACTC	AGTGCGAGTG	ACTGCACTTG	2640
	TCGGGGACCG	CGAGGGCACA	CCTGTCTCCA	TTGTTGTAC	TACGCCGCCT	GAGGCTCCGC	2700
	CAGCCCTGGG	GACGCTTCAC	GTGGTGCAGC	GCGGGGAGCA	CTCGCTGAGG	CTGCGCTGGG	2760
10	AGCCGGTGCC	CAGAGCGCAG	GGCTTCCTTC	TGCACTGGCA	ACCTGAGGGT	GGCCAGGAAC	2820
	AGTCCCGGGT	CCTGGGGCCC	GAGCTCAGCA	GCTATCACCT	GGACGGGCTG	GAGCCAGCGA	2880
	CACAGTACCG	CGTGAGGCTG	AGTGTCTTAG	GGCCGGCTGG	AGAAGGGCCC	TCTGCAGAGG	2940
	TGACTGCGCG	CACTGAGTCA	CCTCGTGTTT	CAAGCATTGA	ACTACGTGTG	GTGGACACCT	3000
	CGATCGACTC	GGTGACTTTG	GCCTGGACTC	CAGTGTCCAG	GGCATCCAGC	TACATCCTAT	3060
15	CCTGGCGGCC	ACTCAGAGGC	CCTGGCCAGG	AAGTGCCTGG	GTCCCCGAG	ACACTTCCAG	3120
	GGATCTCAAG	CTCCAGCGG	GTGACAGGGC	TAGAGCCTGG	CGTCTCTTAC	ATCTTCTCCC	3180
	TGACGCGCTG	CCTGGATGGT	GTGCGGGGTC	CTGAGGCATC	TGTCACACAG	ACGCCAGTGT	3240
	GCCCCCGTGG	CCTGGCGGAT	GTGGTGTTC	TACCACATGC	CACTCAAGAC	AATGCTCACC	3300
	GTGCGGAGGC	TACGAGGAGG	GTCTGGAGC	GTCTGGTGT	GGCACTTGGG	CCTCTTGGGC	3360
20	CACAGGCAGT	TCAGGTTGGC	CTGCTGTCTT	ACAGTCATCG	GCCCTCCCCA	CTGTTCCCAC	3420
	TGAATGGCTC	CCATGACCTT	GGCATTATCT	TGCAAAGGAT	CCGTGACATG	CCCTACATGG	3480
	ACCAAGTGG	GAACAACCTG	GGCAACAGCG	TGGTCACAGC	TCACAGATA	ATGTTGGCAC	3540
	CAGATGCTCC	TGGGCGCCGC	CAGCACGTAC	CAGGGGTGAT	GGTTCTGCTA	GTGGATGAAC	3600
	CCTTGAGAGG	TGACATATTC	AGCCCCATCC	GTGAGGCCCA	GGCTTCTGGG	CTTAATGTGG	3660
25	TGATGTTGGG	AATGGCTGGA	GCGGACCCAG	AGCAGCTGCG	TCGCTTGGCG	CCGGGTATGG	3720
	ACTCTGTCCA	GACCTTCTTC	GCCGTGGATG	ATGGGCCAAG	CCTGGACCAG	GCAGTCAGTG	3780
	GTCTGGCCAC	AGCCCTGTGT	CAGGCATCCT	TCACTACTCA	GCCCCGGCCA	GAGCCCTGCC	3840
	CAGTGTATTG	TCCAAAGGGC	CAGAAGGGGG	AACCTGGAGA	GATGGGCCCT	AGAGGACAAG	3900
	TTGGGGCCTC	TGCGGACCC	GGCCTCCCGG	GCAGGACCCG	TGCTCCCGGC	CCCCAGGGGC	3960
30	CCCCTGGAAG	TGCCACTGCC	AAGGGCGAGA	GGGGCTTCCC	TGGAGCAGAT	GGGCGTCCAG	4020
	GCAGCCCTGG	CCGCGCCGGG	AATCCTGGGA	CCCCTGGAGC	CCCTGGCCTA	AAGGGCTCTC	4080
	CAGGGTTGCC	TGGCCCTCGT	GGGGACCCGG	GAGAGCGAGG	ACCTCGAGGC	CCAAAGGGGG	4140
	AGCCGGGGGC	TCCCGGACAA	GTCTCGGAG	GTGAAGGACC	TGGGCTTCTT	GGGCGGAAAG	4200
	GGGACCCCTG	ACCATCGGGC	CCCCCTGGAG	CTCGTGGACC	ACTGGGGGAC	CCAGGACCCC	4260
35	GTGGCCCCCC	AGGGCTTCTT	AGAACAGCCA	TGAAGGGTGA	CAAAGGCGAT	CGTGGGGAGC	4320
	GGGGTCCCCC	TGGACCAGGT	GAAGGTGGCA	TTGCTCCTGG	GGAGCCTGGG	CTGCCGGGTC	4380
	TTCCCGGAAG	CCCTGGACCC	CAAGGCCCCG	TTGGCCCCCC	TGGAAAGAAA	GGAGAAAAAG	4440
	GTGACTCTGA	GGATGGAGCT	CCAGGCCTCC	CAGGACAACC	TGGGTCTCCG	GGTGAGCAGG	4500
	GCCACGGGG	ACCTCCTGGA	GCTATTGGCC	CCAAAGGTGA	CCGGGGCTTT	CCAGGGCCCC	4560
40	TGGGTGAGGC	TGGAGAGAAG	GGCGAACGTC	GACCCCCAGG	CCCAGCGGGA	TCCCGGGGGC	4620
	TGCCAGGGGT	TGCTGGACGT	CCTGGAGCCA	AGGGTCTCTA	AGGGCCACCA	GGACCCACTG	4680
	GCCGCCAAGG	AGAGAAGGGG	GAGCCTGGTC	GCCCTGGGGA	CCCTGCAGTG	GTGGGACCTG	4740
	CTGTTGCTGG	ACCCAAAGGA	GAAAAGGGAG	ATGTGGGGCC	CGCTGGGCCC	AGAGGAGCTA	4800
	CCGGAGTCCA	AGGGGAACGG	GGCCACCCCG	GCTTGGTTCT	TCCTGGAGAC	CCTGGCCCCA	4860
45	AGGGAGACCC	TGGAGACCGG	GGTCCCATTG	GCCTTACTGG	CAGAGCAGGA	CCCCCAGGTG	4920
	ACTCAGGGCC	TCTTGAGAG	AAGGGAGAGC	TTGGGCGGCC	TGGCCCCCCA	GGACCTGTTG	4980
	GCCCGGAGG	ACGATGGGT	GAAGTTGGAG	AGAAAGGTGA	CGAGGGTCTT	CCGGGTGACC	5040
	CGGGTTTGCC	TGGAAGAGCA	GGCGAGCGTG	GCCTTCGGGG	GGCACCTGGA	GTTCCGGGGC	5100
	CTGTGGGTGA	AAAGGGAGAC	CAGGGAGATC	CTGGAGAGGA	TGGACGAAAT	GGCAGCCCTG	5160
50	GATCATCTGG	ACCAAGGGT	GACCGTGGGG	AGCCGGGTCC	CCCAGGACCC	CCGGGACGGC	5220
	TGGTAGACAC	AGGACCTGGA	GCCAGAGAGA	AGGGAGAGCC	TGGGGACCGC	GGACAAGAGG	5280
	GTCTCTGAGG	GCCCAAGGGT	GATCCTGGCC	TCCCTGGAGC	CCCTGGGGAA	AGGGGCATTG	5340
	AAGGGTTTCG	GGGACCCCCA	GGGACCCAGG	TGTCCGAGGC	CCAGCAGGAG	5400	
	AAAAGGGTGA	CCGGGGTCCC	CCTGGGCTGG	ATGGCCGGAG	CGGACTGGAT	GGGAAACCAG	5460
55	GAGCCGCTGG	GCCCTCTGGG	CCGAATGGTG	CTGCAGGCAA	AGCTGGGGAC	CCAGGGAGAG	5520
	ACGGGCTTCC	AGGCCTCCGT	GGAGAACAAG	GCCTCCCTGG	CCCCTCTGGT	CCCCCTGGAT	5580
	TACCGGGAAG	GCCAGGCGAG	GATGGGAAAC	CTGGCCTGAA	TGGAAAAAAC	GGAGAACCCTG	5640
	GGGACCCCTG	AGAAGACGGG	AGGAAGGGAG	AGAAAGGAGA	TTCAGGCGCC	TCTGGGAGAG	5700
	AAGGTCTGTA	TGGCCCCAAG	GGTGAGCGTG	GAGCTCTCTG	TATCCTTGGG	CCCCAGGGGC	5760
60	CTCCAGGCCCT	CCAGGGGCCA	GTGGGCCCTC	CTGGCCAGGG	TTTTCCTGGT	GTCCAGGAG	5820
	GCACGGGCCC	CAAGGGTGAC	CGTGGGGAGA	CTGGATCCAA	AGGGGAGCAG	GGCCTCCCTG	5880
	GAGAGCGTGG	CCTGCGAGGA	GAGCCTGGAA	GTGTGCCGAA	TGTGGATCGG	TTGCTGGAAA	5940
	CTGCTGGCAT	CAAGGCATCT	GCCCTGCGGG	AGATCGTGGA	GACCTGGGAT	GAGAGCTCTG	6000
	GTAGCTTCTT	GCCTGTGCCC	GAACGGCGTC	GAGGCCCCAA	GGGGGACTCA	GGCGAACAGG	6060
65	GCCCCCAGG	CAAGGAGGGC	CCCATCGGCT	TTCTTGAGGA	ACGCGGGCTG	AAGGGCGACC	6120
	GTGGAGACCC	TGGCCCTCAG	GGGCCACCTG	GTCTGGCCCT	TGGGGAGAGG	GGCCCCCCCG	6180
	GGCCTTCCGG	CCTTGCCGGG	GAGCCTGGAA	AGCCTGGTAT	TCCCGGGCTC	CCAGGCAGGG	6240
	CTGGGGGTGT	GGGAGAGGCA	GGAAGGCCAG	GAGAGAGGGG	AGAACGGGGA	GAGAAAGGAG	6300
	AACGTGGAGA	ACAGGGCAGA	GATGGCCCTC	CTGGACTCCC	TGGAAACCCCT	GGGCCCCCCG	6360
70	GACCCCTGG	CCCCAAGGTG	TCTGTGGATG	AGCCAGGTCC	TGGACTCTCT	GGAGAACAGG	6420
	GACCCCTGG	ACTCAAGGGT	GCTAAGGGGG	AGCCGGGCGG	CAATGGTGAC	CAAGGTCCCA	6480
	AAGGAGACAG	GGGTGTGCCA	GGACCTCAAG	GAGACCCGGG	AGAGCCTGGA	CCGAGGGGTC	6540
	AGGACGGCAA	CCCGGGTCTA	CCAGGAGAGC	GTGGTATGGC	TGGGCCTGAA	GGGAAGCCGG	6600
	GTCTGCAGGG	TCCAAGAGGC	CCCCCTGGCC	CAGTGGGTGG	TCATGGAGAC	CCTGGACCAC	6660
75	CTGGTGCCCC	GGGTCTTGCT	GGCCCTGCAG	GACCCCAAGG	ACCTTCTGGC	CTGAAGGGGG	6720
	AGCCTGGAGA	GACAGGACCT	CCAGGACGGG	GCCTGACTGG	ACCTACTGGA	GCTGTGGGAC	6780
	TTCTTGAGCC	CCCCGGCCCT	TCAGGCCTTG	TGGGTCCACA	GGGGTCTCCA	GGTTTGCTTG	6840
	GACAAGTGGG	GGAGACAGGG	AAGCCGGGAG	CCCCAGGTCC	AGATGGTGCC	AGTGGAAAAG	6900
	ATGGAGACAG	AGGGAGCCCT	GGTGTGCCAG	GGTCAACCAG	TCTGCCTGGC	CCTGTGCGAC	6960
80	CTAAAGGAGA	ACCTGGCCCC	ACGGGGGCCC	CTGGACAGGC	TGTGGTCCGG	CTCCCTGGAG	7020
	CAAAGGGAGA	GAAGGGAGCC	CCTGGAGGCC	TTGCTGGAGA	CCTGGTGGGT	GAGCCGGGAG	7080
	CCAAAGGTGA	CCGAGGACTG	CCAGGGCCGC	GAGGCGAGAA	GGGTGAAGCT	GGCCGTGCAG	7140
	GGGAGCCCGG	AGACCTGGGG	GAAGATGGTC	AGAAAGGGGC	TCCAGGACCC	AAAGGTTTCA	7200
	AGGGTGACCC	AGGAGTCGGG	GTCCCGGGCT	CCCCTGGGCC	TCCTGGCCCT	CCAGGTGTGA	7260
85	AGGGAGATCT	GGGCTTCCCT	GGCCTGCCCC	GTGCTCCTGG	TGTTGTTGGG	TTCCCGGGTC	7320
	AGACAGGCCC	TCGAGGAGAG	ATGGGTGAGC	CAGGCCCTAG	TGGAGAGCGG	GGTCTGGCAG	7380
	GCCCCCAGG	GAGAGAAGGA	ATCCAGGAC	CCCTGGGGCC	ACCTGGACCA	CCGGGGTCAG	7440
	TGGGACCACC	TGGGGCTCTT	GGACTCAAAG	GAGACAAGGG	AGACCTTGA	GTAGGGCTGC	7500



	CTGGGCCCCG	AGGCGAGCGT	GGGGAGCCAG	GCATCCGGGG	TGAAGATGGC	CGCCCCGGCC	7560
	AGGAGGGACC	CCGAGGACTC	ACGGGGCCCC	CTGGCAGCAG	GGGAGAGCGT	GGGGAGAAGG	7620
	GTGATGTTGG	GAGTGCAGGA	CTAAAGGGTG	ACAAGGGAGA	CTCAGCTGTG	ATCCTGGGGC	7680
5	CTCCAGGCCC	ACGGGGTGCC	AAGGGGGACA	TGGGTGAACG	AGGGCCTCGG	GGCTTGATG	7740
	GTGACAAAGG	ACCTCGGGGA	GACAATGGGG	ACCCTGGTGA	CAAGGGCAGC	AAGGGAGAGC	7800
	CTGGTGACAA	GGGCTCAGCC	GGGTTGCCAG	GA CTGCGTGG	ACTCCTGGGA	CCCCAGGGTC	7860
	AACCTGGTGC	AGCAGGGATC	CCTGGTGACC	CGGGATCCCC	AGGAAAGGAT	GGAGTGCCTG	7920
	GTATCCGAGG	AGAAAAAGGA	GATGTTGGCT	TCATGGGTCC	CCGGGGCCTC	AAGGGTGAAC	7980
10	GGGGAGTGAA	GGGAGCCTGT	GGCCTTGATG	GAGAGAAGGG	AGACAAGGGA	GAAGCTGGTC	8040
	CCCCAGGCCC	CCCCGGGCTG	GCAGGACACA	AAGGAGAGAT	GGGGGAGCCT	GGTGTGCCGG	8100
	GCCAGTCGGG	GGCCCCCTGG	AAGGAGGGCC	TGATCGGTCC	CAAGGGTGAC	CGAGGCTTTG	8160
	ACGGGCAGCC	AGGCCCAAG	GGTGACCAGG	GCGAGAAAGG	GGAGCGGGGA	ACCCAGGAA	8220
	TTGGGGGCTT	CCCAGGCCCC	AGTGGAATG	ATGGCTCTGC	TGGTCCCCCA	GGGCCACCTG	8280
	GCACTGTTGG	TCCAGAGGCG	CCCGAAGGAG	TTCAGGGCCA	GAAGGGTGAG	CGAGGTCCCC	8340
15	CCGGAGAGAG	AGTGGTGGGG	GCTCCTGGGG	TCCCTGGAGC	TCCTGGCGAG	AGAGGGGAGC	8400
	AGGGGCGGCC	AGGGCCTGCC	GGTCTCGAG	GCGAGAAGGG	AGAAGCTGCA	CTGACGGAGG	8460
	ATGACATCCG	GGGCTTTGTG	CGCCAAGAGA	TGAGTCAGCA	CTGTGCCTGC	CAGGGCCAGT	8520
	TCATCGCATC	TGGATCACGA	CCCCTCCCTA	GTTATGCTGC	AGACACTGCC	GGCTCCCAGC	8580
20	TCCATGCTGT	GCCTGTGCTC	CGCGTCTCTC	ATGCAGAGGA	GGAAGAGCGG	GTACCCCTTG	8640
	AGGATGATGA	GTACTCTGAA	TACTCCGAGT	ATTCTGTGGA	GGAGTACCAG	GACCCCTGAAG	8700
	CTCCTTGGGA	TAGTGATGAC	CCCTGTTCCT	TGCCACTGGA	TGAGGGCTCC	TGCACTGCCT	8760
	ACACCTTGCG	CTGGTACCAT	CGGGCTGTGA	CAGGCAGCAC	AGAGGCCTGT	CACCCCTTTG	8820
	TCTATGGTGG	CTGTGGAGGG	AATGCCAAC	GTTTTGGGAC	CCGTGAGGCC	TGCGAGCGCC	8880
25	GCTGCCACCC	CCGGGTGGTC	CAGAGCCAGG	GGACAGGTAC	TGCCCAGGAC	TGAGGCCCAG	8940
	ATAATGAGCT	GAGATTGAGC	ATCCCCTGGA	GGAGTCGGGG	TCTCAGCAGA	ACCCCACTGT	9000
	CCCTCCCCTT	GGTGCTAGAG	GCTTGTGTGT	ACGTGAGCGT	GCGAGTGAC	GTCCGTTATT	9060
	TCAGTGACTT	GGTCCCGTGG	GTCTAGCCTT	CCCCCTGTG	GACAAACCCC	CATTGTGGCT	9120
	CCTGCCACCC	TGGCAGATGA	CTCACTGTGG	GGGGGTGGCT	GTGGGCAGTG	AGCGGATGTG	9180
30	ACTGGCGTCT	GACCCGCCCC	TTGACCCAAG	CCTGTGATGA	CATGGTGCTG	ATTCTGGGGG	9240
	GCATTAAAGC	TGCTGTTTTA	AAAGGCAAAA	AA			

Seq ID NO: 63 Protein sequence:  
Protein Accession #: NP\_000085.1

35	1	11	21	31	41	51	
	MTLRLIVAAL	CAGILAEAPR	VRAQHRERVT	CTRLAADIIV	FLLDGSSSIG	RSNFREVRSF	60
	LEGLVLPFSG	AASAQGVRF	TVQYSDDPRT	EFGLDALGSG	GDVIRAIREL	SYKGGNTRTG	120
40	AAILHVADHV	FLPQLARPGV	PKVCILITDG	KSQDLVDTA	QRLKGQVKL	FAVGKINADP	180
	EELKRVASQP	TSDFFFVND	FSILRTLPL	VSRRVCTTAG	GVPVTRPPDD	STSAPRDLVL	240
	SEPSSQSLRV	QWTAASGPVT	GYKVQYTLPT	GLGQPLPSE	QEVNVPAGET	SVRLRGLRPL	300
	TEYQVTIVIAL	YANSIGEAVS	GTARTTALEG	PELTIQNTTA	HSLLVAVRSV	PGATGYRVTW	360
	RVLSSGGPTQ	QELGPGQGSV	LLRDLEPGTD	YEVTVSTLFG	RSVGPATSLM	ARTDASVEQT	420
45	LRPVILGPTS	ILLSWNILVPE	ARGYRLEWRR	ETGLEPPQKV	VLPSDVTRYQ	LDGLQPGETEY	480
	RLTLTYLLEG	HEVATPATV	PTGPELFPVS	VTDLQATELP	GQRVRVSWSP	VPGATQYRII	540
	VRSTQGVERT	LVLPQSQTAF	DLDDVQAGLS	YTVRV SARVG	PREGSASVLT	VRREPETPLA	600
	VPGLRVVVS	ATRVVAVWGP	VPGASGFRIS	WSTGSGPESS	QTLPPDSTAT	DITGLQPGTT	660
	YQVAVSVLRG	REEGPAAVIV	ARTDPLGPVR	TVHVTQASS	SVTITWTRVP	GATGYRVSWH	720
50	SAHGPEKSQL	VSGEATVAEL	DGLEPDTEYT	VHVRHVAGV	DGPPASVVVR	TAPEPVGRVS	780
	RLQILNASSD	VLRLITWVGT	GATAYRLAWG	RSEGGPMRHO	ILPGNTDSAE	IRGLEGGVSY	840
	SVRVTAIVGD	REGTPVSIVV	TTPPEAPPAL	GTLVHVQRGE	HSLRLRWEVP	PRAQGFLLHW	900
	QPEGGEQESR	VLGPELSSYH	LDGLEPATYQ	RVLRLSVLGA	GEGPSAEVTA	RTESPRVPSI	960
	ELRVVDTSID	SVTLAWTPVS	RASSYILSWR	PLRGPQGVEP	GSPQTLPGIS	SSQRTGLEP	1020
55	GVSIFYSLTP	VLDGVRGPEA	SVTQTPVCPR	GLADVFLPH	ATQDNAHRAE	ATRRVLERLV	1080
	LALGPLGPQA	VQVGLLSYSH	RPSPLFPLNG	SHDLGIILQR	IRDMPYMDPS	GNNLGTAVVT	1140
	AHRYMLAPDA	PGRRQHVPGV	MVLLVDEPLR	GDIFSPIREA	QASGLNVVML	GMAGADPEQL	1200
	RRLAPGMDSV	QTFFAVDDGP	SLDQAVSGLA	TALCQASFTT	QPRPEPCPVY	CPKGQKGEPE	1260
	EMGLRGQVGP	PGDPGLPGRT	GAPGPQGPFG	SATAGGERGF	PGADGRPGSP	GRAGNPCTPG	1320
60	APGLKGSPL	PGPRGDPGER	PGPRGDPGER	APGQVIGGEG	PGLPGRKGD	GPSGPPGPRG	1380
	PLGDPGPRGP	PGLPGTAMKG	DKGDRGERGP	PGPGEGLIAP	GEPGLPGLPG	SPGPQGPVGP	1440
	PGKKGEKGD	EDGAPGLPGQ	PGSPGEQGP	GPPGAIGPKG	DRGFPGLGE	AGEKGERGPP	1500
	GPAGSRGLPG	VAGRPAGAKP	EGPPGPTRGQ	GEKGPGRPG	DPAVVGPAVA	GPKGEKGDVG	1560
	PAGPRGATGV	QGERGPPGLV	LPGDPGPKGD	PGDRGPGLT	GRAGPPGDSG	PPGEKGDPR	1620
	PGPPGPVGP	GRDGEVGEK	DEGPPGDPGL	PGKAGERGLR	GAPGVRGPVG	EKGDQGDPE	1680
65	DGRNGSPGSS	GPKGDRGEPG	PPGPPGRLVD	TGPGAREKGE	PGDRGQEGPR	GPKGDPGLPG	1740
	APGERGIEGF	RGPDPGQD	GVRGPAGEK	DRGPPGLDGR	SGLDGKPGAA	GPSGPNGAAG	1800
	KAGDPGRDGL	PGLRGEQGLP	GPSGPPGLPG	KPGEDGKPL	NGKNGEPGD	GEDGRKGEK	1860
	DSGASGREGR	DGPKGERGAP	GILGPQGPFG	LPGPVGPFGQ	GFPVGPFGTG	PKGDRGETGS	1920
70	KGEQGLPGER	GLRGEFGSVP	NVDRLETAG	IKASALREIV	ETWDESSGSF	LPVPERRRGP	1980
	KGDSGEQGP	GKEGPIGFPG	ERGLKDRGD	PGPQGPGLA	LGERGPPGPS	GLAGEPGKPG	2040
	IPGLPGRAGG	VGEAGRPGER	GERGEKGERG	EQGRDGPGL	PCTPGPPGP	GPKVSVDEPG	2100
	PGLSGEQGP	GLKGAKGEPG	SNGDQGPKG	RGVPGIKGDR	GEPGPRGQDG	NPGLPGERGM	2160
	AGPEGKPLQ	GPRGPPGPVG	GHGDPGPPGA	PGLAGPAGPQ	GPSGLKGEPE	ETGPPGRGLT	2220
75	GPTGAVGLPG	PPGPSGLVGP	QGSPLPGQV	GETGKPGAPG	RDGASGKGD	RGSPGVPGSP	2280
	GLPGFVGPKG	EPGPTGAPGQ	AVVGLPGAKG	EKGAPGGLAG	DLVGEPGAKG	DRGLPGPRGE	2340
	KGEAGRAGEP	GDPGEDGQKG	APGPKGFKGD	PGVGVPGSPG	PPGPPGVKGD	LGLPGLPGAP	2400
	GVVGFPGQTG	PRGEMGQPGP	SGERGLAGPG	GREGIPGLG	PPGPPGSGVP	PGASGLKGD	2460
	GDPGVGLPGP	RGERGEPGIR	GEDGRPGQEG	PRGLTGPPGS	RGERGEKGDV	GSAGLKGDG	2520
80	DSAVILGPPG	PRGAKGDMGE	RGRGLDGDG	GPRGNDGDPG	DKGSKGEPGD	KGSAGLPGLR	2580
	GLLGPQGP	AAGIPGDPGS	PGKDVGPGR	GEKGDVGMFG	PRGLKGERGV	KGACGLDGEK	2640
	GDKGEAGPPG	RPGLAGHKGE	MGEPPGVPGS	GAPGKEGLIG	PKGDRGFDGQ	PGPKGDQGEK	2700
	GERGTPGIGG	FPGPSGNDGS	AGPPGPPGVS	GPRGPEGLQG	QKGERGPPGE	RVVGAPGVPG	2760
	APGERGEQGR	PGPAGPRGEK	GEAALTEDDI	RGFVRQEVSQ	HCACQGFIA	SGSRPLPSYA	2820
85	ADTAGSQLHA	VPVLRVSHAE	EEERVPPEDD	EYSEYSEYSV	EEYQDPEAPW	DSDDPCSLPL	2880
	DEGSCTAYTL	RWYHRAVTGS	TEACHPFVYG	GCGGNANRFG	TREACERRCP	PRVVQSQGTG	2940
	TAQD						



**PCT/US02/12476**

Seq ID NO: 65 Protein sequence:  
Protein Accession #: NP\_008876

Seq ID NO: 66 DNA sequence  
Nucleic Acid Accession #: NM\_005629.1  
Coding sequence: 639-2546

213



CTTCTGTGTA GCAGCTTTAA CCCACGTTTG TCTGTCACGT CCAGTCCCGA GACGGCTGAG 3780  
 TGACCCCAAG AAAGGCTTCC CCGACACCCA GACAGAGGCT GCAGGGCTGG GGCTGGGTGA 3840  
 GGGTGGCGGG CCTGCGGGGA CATTCTACTG TGCTAAAAAG CCACTGCAGA CATAGCAATA 3900  
 AAAACATGTC ATTTTCC

Seq ID NO: 67 Protein sequence:  
 Protein Accession #: NP\_005620.1

1 11 21 31 41 51  
 | | | | |  
 MAKSAENGI YSVSGDEKKG PLIAPGPDGA PAKGDGPVGL GTPGGRLAVP PRETWTRQMD 60  
 FIMSCVGFV GLGNVWRFPY LCKYKNGGCVF LIPYVLIALV GGIPIFFLEI SLGQFMKAGS 120  
 INVWNICPLF KGLGYASPMI VFYCNYYIIM VLAWGFYYLV KSFTTTLPLWA TCIGHTWNTPD 180  
 CVEIFRHEHC ANASLANLTC DQLADRRSPV IEFWENKVLK LSGGLEVPFA LNWEVTLCLL 240  
 ACWVLVYFCV WKGVKSTGKI VYFTATFPYV VLVLVLRGV LLPGALDGI IYYLKPDWSKL 300  
 GSPQVWIDAG TQIFFSYAIG LGALTALGSY NFRNNNCYKD AILALINSI TSFFAGFVVF 360  
 SILGFMAAEQ GVHISKVAES GPGLAFIAYP RAVTLMPVAP LWAAALFFFML LLLGLDSQFV 420  
 GVEGFTITGL DLLPASYYFR FQREISVALC CALCFVIDLS MVTGGMVVF QLFDYYSASG 480  
 TLLWQAFWE CVVVAWVYGA DRFMDDIACM IGYRCPWMK WCWSFFTPLV CMGIFIFNVV 540  
 YYPELVYNN YVYPWWEAM GWAFALSSML CVPLHLLGCL LRAGKTMAER WQHLTQPIWG 600  
 LHHLEYRAQD ADVRGLTTLT PVSESSKVVV VESVM

Seq ID NO: 68 DNA sequence  
 Nucleic Acid Accession #: NM\_021953.1  
 Coding sequence: 178-2469

1 11 21 31 41 51  
 | | | | |  
 GGCACGAGGG GGACCCGGCC GGTCCGGCGC GAGCCCCCGT CCGGGGCCCT GGCTCGGCC 60  
 CCAGGTTGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG GCGGCGACTG 120  
 CAGTCTGGAG GGTCCACACT TGTGATTCTC AATGGAGAGT GAAAACGCAG ATTCATAATG 180  
 AAAGCTAGCC CCCGTCGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCCC TCCTGTTCAA 240  
 AATGCCCCAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCCCTGCCCA ACAGGAGTCT 300  
 AATCAAGCAG AGGCCTCCAA GGAAGTGGCG GAGTCCAACCT CTTGCAAGTT TCCAGCTGGG 360  
 ATCAAGATTA TTAACCAACC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCAACAAT 420  
 GCTAATATTC ACAGCATCAT CACAGCACTG ACTGCCAAGG GAAAAGAGAG TGGCAGTAGT 480  
 GGGCCCAACA AATTCATCCT CATCAGCTGT GGGGGAGCCC CAACTCAGCC TCCAGGACTC 540  
 CGGCCCTCAA CCCAAACCAG CTATGATGCC AAAAGGACAG AAGTGACCCT GGAGACCTTG 600  
 GGACCAAAAC CTGCGCTAGC GATGTGAAT CTTCCCTAGC CACTGGAGC CCTTTGCGAG 660  
 CAGAAACGGG AGACCTGTGC AGATGGTGAG GCAGCAGGCT GCACTATCAA CAATAGCCTA 720  
 TCCAACATCC AGTGGCTTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780  
 CAAGAGATGG AGGAAAAGGA GAATTGTAC CTGGAGCAGC GACAGGTTAA GGTGAGGAG 840  
 CCTTCGAGAC CATCAGCGTC CTGGCAGAAC TCTGTGTCTG AGCGGCCACC CTACTCTTAC 900  
 ATGGCCATGA TACAATTCGC CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960  
 ATCTATACGT GGATTGAGGA CCAATTTCCC TACTTTAAGC ACATTGCCAA GCCAGGCTGG 1020  
 AAGAATCCCA TCCGCCACAA CCTTTCCCTG CACGACATGT TTGTCCGGGA GACGTCTGCC 1080  
 AATGGCAAGG TCTCCTTCTG GACCATTAC CCCAGTGCCA ACCGCTACTT GACATTGGAC 1140  
 CAGGTGTTTA AGCCACTGGA CCCAGGGTCT CCACAATTGC CCGAGCACTT GGAATCACAG 1200  
 CAGAAACGAC CGAATCCAGA GCTCCGCGCG AACATGACCA TCAAAACCGA ACTCCCCCTG 1260  
 GGCGCACGGC GGAAGATGAA GCCACTGCTA CCACGGGTCA GCTCATACCT GGTACCTATC 1320  
 CAGTTCCTCG TGAACAGTC ACTGGTGTGG GACCCCTCGG TGAAGGTGCC ATTGCCCTTG 1380  
 GCGGCTTCCC TCATGAGCTC AGAGCTTGCC CGCCATAGCA AGCGAGTCCG CATTGCCCTC 1440  
 AAGGTGCTGC TAGCTGAGGA GGGGATAGCT CCTCTTTCTT CTGCAGGACC AGGGAAAGAG 1500  
 GAGAACTCC TGTGAGGAG AGGGTTTTCT CCTTTGCTTC CAGTTCAGAC TATCAAGGAG 1560  
 GAAGAAATCC AGCCTGGGGA GGAAATGCCA CACTTAGCGA GACCCATCAA AGTGGAGAGC 1620  
 CCTCCCTTGG AAGAGTGGCC CTCCCCGGCC CCATCTTTCA AAGAGGAATC ATCTCACTCC 1680  
 TGGGAGGATT CGTCCCAATC TCCCACCCCA AGACCCAAGA AGTCCTACAG TGGGCTTAGG 1740  
 TCCCCAACC GGTGTGTCTC GGAAATGCTT GTGATTCAAC ACAGGGAGAG GAGGGAGAGG 1800  
 AGCCGCTCTC GGAGGAAACA GCATCTACTG CCTCCCTGTG TGGATGAGCC GGAGCTGCTC 1860  
 TTCTCAGAGG GGCCAGTAC TTCCCGCTGG GCCGAGAGC TCCCGTTCCC AGCAGACTCC 1920  
 TCTGACCTCG CCTCCAGCT CAGCTACTCC CAGGAAGTGG GAGGACCTTT TAAGACACCC 1980  
 ATTAAGGAAA CGCTGCCCAT CTCCTCCACC CCGAGCAAAT CTGTCTCTCC CAGAACCCCT 2040  
 GAATCCTGGA GGCTCAGCC CCCAGCCAAA GTAGGGGAGC TGGATTTCAG CCCAGTACAA 2100  
 ACCTCCAGG GTGCTCTGA CCCCCTGCTT GACCCCTGGG GGCTGATGGA TCTCAGCACC 2160  
 ACTCCCTTGC AAAGTGCTCC CCCCCTTGAA TCACCGCAAA GGCTCCTCAG TTCAGAACCC 2220  
 TTAGACCTCA TCTCCGTCCC CTTTGGCAAC TCTTCTCCCT CAGATATAGA CGTCCCCAAG 2280  
 CCAGGCTCCC CGGAGCCACA GGTGCTGAG CTTGAGGCA ATCGTTCTCT GACAGAAGGC 2340  
 CTGGTCTCTG ACACAATGAA TGACAGCCTC AGCAAGATCC TGCTGGACAT CAGCTTTCTC 2400  
 GGCTGGACG AGGACCCACT GGGCCCTGAC AACATCAACT GGTCCAGTT TATTCCTGAG 2460  
 CTACAGTAGA GCCCTGCCCT TGCCCTGTG CTCAAGCTGT CCACCATCCC GGGCACTCCA 2520  
 AGGCTCAGTG CACCCCAAGC CTCTGAGTGA GGACAGCAGG CAGGGACTGT TCTGCTCCTC 2580  
 ATAGCTCCCT GCTGCCTGAT TATGCAAAAG TAGCAGTCAC ACCCTAGCCA CTGCTGGGAC 2640  
 CTTGTGTTCC CCAAGAGTAT CTGATTCTCT TGCTGTCCCT GCCAGGAGCT GAAGGGTGGG 2700  
 AACAAACAAAG GCAATGGTGA AAAGAGATTA GGAACCCCCC AGCCTGTTTC CATTCTCTGC 2760  
 CCAGCAGTCT CTTACCTTCC CTGATCTTTG CAGGGTGGTC CGTGTAATA GTATAAATTC 2820  
 TCCAAATTAT CCTCTAATTA TAAATGTAAG CTATTTCCT TAGATCATT TCCAGAGACT 2880  
 GCCAGAGGT GGTGAGTAT ACCTGGGGTT TCAATTGACT TCTGTTCTTT ACCTTTTAGT 2940  
 TTGATAGAAG GGAAGACCTG CAGTGCACGG TTTCTTCCAG GCTGAGGTAC CTGGATCTTG 3000  
 GGTCTTTCAC TGCAGGGACC CAGACAAGTG GATCTGCTTG CCAGAGTCTT TTTTGCCCTC 3060  
 CCCTGCCACC TCCCGTGTG TCCAAGTCAG CTTTCCCTGCA AGAAGAAATC CTGGTTAAAA 3120  
 AAGTCTTTG TATTGGGTCA GGAGTTGAAT TTGGGGTGGG AGGATGGATG CAACTGAAGC 3180  
 AGAGTGTGGG TGCCAGATG TGCGCTATTA GATGTTCTC TGATAATGTC CCCAATCATA 3240  
 CCAGGGAGAC TGGCATTGAC GAGAATTGAG GTGGAGGCTT GAGAAGCCG AAAGGGCCCC 3300  
 TGACCTGCCCT GGCTTCTTGA GCTTGGCCCT CAGCTTTGCA AAGAGCCACC CTAGGCCCCA 3360  
 GCTGACCGCA TGGGTGTGAG CCAGCTTGAG AACACTAACT ACTCAATAAA AGCGAAGGTG 3420  
 GACCNAAAAA AAAAAAAAAA AAAA



Seq ID NO: 69 Protein sequence:  
Protein Accession #: NP\_068772.1

5	1	11	21	31	41	51	
	MKASPRRPLI	LKRRRLPLPV	QNAPSETSEE	EPKRSPAQQE	SNQAEASKEV	AESNSCKFPA	60
	GIKIINHPTM	PNTQVVAIPN	NANIHSIITA	LTAKGKESGS	SGPNKFILIS	CGGAPTQPPG	120
	LRPQTQTSYD	AKRTEVTLET	LGPKPAARDV	NLPRPPGALC	EOKRETCADG	EAAGCTINNS	180
10	LSNIQWLKRM	SSDGLGSRSI	KQEMEKEKENC	HLEQRQVKVE	EPSRPSASWQ	NSVSRPPYS	240
	YMAMIQFAIN	STERKRMTLK	DIYTWIEDHF	PYFKHIAKPG	WKNSIRHNLS	LHDMFVRETS	300
	ANGKVSFWTI	HPSANRYLTL	DQVFKPLDPG	SPQLPEHLES	QOKRPNPELR	RNMTIKTELP	360
	LGARRKMKPL	LPRVSSYLVP	IQFPVNQSLV	LQPSVKVPLP	LAASLMSEEL	ARHSKRVRIA	420
	PKVLLAEEGI	APLSSAGPGK	EEKLLFGEFG	SPLLVPQTIK	EEEIQPGGEM	PHLARPIKVE	480
15	SPPLEEWSP	APSFKEESSH	SWEDSSQSPT	PRPKKSYSGL	RSPTRCVSEM	LVIQHREERE	540
	RSRSRRKQHL	LPCCVDEPEL	LFSEGPSTSR	WAAELPFPAD	SSDPASQLSY	SQEVGGPFKT	600
	PIKETLPISS	TPSKSVLPRT	PESWRLTPPA	KVGGLDFSPV	QTSQGASDPL	PDPLGLMDLS	660
	TTPLQSAPPL	ESPQRLLSSE	PLDLISVPFG	NSSPSDIDVP	KPGSPPEQVS	GLAANRSLTE	720
20	GLVLDTMND	LSKILLDISF	PGLDEDPLGP	DNINWSQFIP	ELQ		

Seq ID NO: 70 DNA sequence  
Nucleic Acid Accession #: BC006529.1  
Coding sequence: 178-2424

25	1	11	21	31	41	51	
	GGCACGAGGG	GGACCCGGCC	GGTCCGGCGC	GAGCCCCCGT	CCGGGGCCCT	GGCTCGGCC	60
	CCAGGTGGA	GGAGCCCGGA	GCCCGCCTTC	GGAGCTACGG	CCTAACGGCG	GCGGCGACTG	120
30	CAGTCTGGAG	GGTCCACACT	TGTGATTCTC	AATGGAGAGT	GAAAACGCAG	ATTCTAATG	180
	AAAACGTAGC	CCCGTCGGCC	ACTGATTCTC	AAAAGACGGA	GGCTGCCCC	TCCTGTTCAA	240
	AATGCCCAA	GTGAAACATC	AGAGGAGGAA	CCTAAGAGAT	CCCCTGCCCA	ACAGGAGTCT	300
	AATCAAGCAG	AGGCTTCAA	GGAAGTGGCA	GAGTCCAAT	CTTGCAAGTT	TCCAGCTGGG	360
	ATCAAGATTA	TTAACCACCC	CACCATGCC	AACACGCAAG	TAGTGGCCAT	CCCCAACAA	420
35	GCTAATATTC	ACAGCATCAT	CACAGCACTG	ACTGCCAAGG	GAAAAGAGAG	TGGCAGTAGT	480
	GGGCCCCAACA	AATTCATCCT	CATCAGCTGT	GGGGGAGCCC	CAACTCAGCC	TCCAGGACTC	540
	CGGCCTCAA	CCCAAACCAG	CTATGATGCC	AAAAGGACAG	AAGTGACCCT	GGAGACCTTG	600
	GGACCAAAAC	CTGCAGCTAG	GGATGTGAAT	CTTCCTAGAC	CACCTGGAGC	CCTTTGCGAG	660
	CAGAAACGGG	AGACCTGTGC	AGATGGTGAG	GCAGCAGGCT	GCACTATCAA	CAATAGCCTA	720
40	TCCAACATCC	AGTGGCTTCG	AAAGATGAGT	TCTGATGGAC	TGGGCTCCCG	CAGCATCAAG	780
	CAAGAGATGG	AGGAAAAGGA	GAATTGTGAC	CTGGAGCAGC	GACAGGTTAA	GGTTGAGGAG	840
	CCTTCGAGAC	CATCAGCGTC	CTGGCAGAAC	TCTGTGTCTG	AGCGGCCACC	CTACTCTTAC	900
	ATGGCCATGA	TACAATTCGC	CATCAACAGC	ACTGAGAGGA	AGCGCATGAC	TTTGAAAGAC	960
45	ATCTATACCT	GGATTGAGGA	CCACTTTCCC	TACTTTAAGC	ACATTGCCAA	GCCAGGCTGG	1020
	AAGAACTCCA	TCCGCCACAA	CCTTTCCCTG	CACGACATGT	TTGTCCGGGA	GACGTCTGCC	1080
	AATGGCAAGG	TCTCCTTCTG	GACCATTCAC	CCCAGTGCCA	ACCGTACTT	GACATTGGAC	1140
	CAGGTGTTTA	AGCAGCAGAA	ACGACCGAAT	CCAGAGCTCC	GCCGGAACAT	GACCATCAAA	1200
	ACCGAACTCC	CCCTGGGCGC	ACGGCGGAAG	ATGAAGCCAC	TGCTACCACG	GGTCAGCTCA	1260
	TACCTGGTAC	CTATCCAGTT	CCCGGTGAAC	CAGTCACTGG	TGTTGCAGCC	CTCGGTGAAG	1320
50	GTGCCATTGC	CCCTGGCGGC	TTCCCTCATG	AGCTCAGAGC	TTGCCCGCCA	TAGCAAGCGA	1380
	GTCCGCATTG	CCCCCAAGGT	GCTGCTAGCT	GAGGAGGGGA	TAGCTCCTCT	TTCTTCTGCA	1440
	GGACCAGGGA	AAGAGGAGAA	ACTCCTGTTT	GGAGAAGGGT	TTTCTCCTTT	GCTTCCAGTT	1500
	CAGACTATCA	AGGAGGAAGA	AATCCAGCCT	GGGAGGAAAG	TGCCACACTT	AGCGAGACCC	1560
	ATCAAAGTGG	AGAGCCCTCC	CTTGGAAGAG	TGGCCCTCCC	CGGCCCATC	TTTCAAAGAG	1620
55	GAATCATCTC	ACTCCTGGGA	GGATTCTGTC	CAATCTCCCA	CCCCAAGACC	CAAGAAGTCC	1680
	TACAGTGGGC	TTAGGTCCCC	AACCCGGTGT	GTCTCGGAAA	TGCTTGTGAT	TCAACACAGG	1740
	GAGAGGAGGG	AGAGGAGCCG	GTCTCGGAGG	AAACAGCATC	TACTGCCTCC	CTGTGTGGAT	1800
	GAGCCGGAGC	TGCTCTTCTC	AGAGGGGCCC	AGTACTTCCC	GCTGGGCGC	AGAGCTCCCG	1860
	TTCCAGCAGC	ACTCCTCTGA	CCCTGCCTCC	CAGTCAAGCT	ACTCCAGGA	AGTGGGAGGA	1920
60	CCTTTAAAGA	CACCCATTAA	GGAAACGCTG	CCCATCTCCT	CCACCCCGAG	CAAATCTGTC	1980
	CTCCCCAGAA	CCCCTGAATC	CTGGAGGCTC	ACGCCCCCAG	CCAAAGTAGG	GGGACTGGAT	2040
	TTCAGCCCAG	TACAAACCCC	CCAGGGTGCC	TCTGACCCCT	TGCCTGACCC	CCTGGGGCTG	2100
	ATGGATCTCA	GCACCACTCC	CTTGCAAGT	GCTCCCCCCC	TTGAATCACC	GCAAAGGCTC	2160
	CTCAGTTCAG	AACCTTAGA	CCTCATCTCC	GTCCCCCTTG	GCAACTCTTC	TCCCTCAGAT	2220
65	ATAGACGTCC	CCAAGCCAGG	CTCCCCGGAG	CCACAGGTTT	CTGGCCTTGC	AGCCAATCGT	2280
	TCTCTGACAG	AAGGCCTGGT	CCTGGACACA	ATGAATGACA	GCCTCAGCAA	GATCCTGCTG	2340
	GACATCAGCT	TTCTGGCCT	GGACGAGGAC	CCACTGGGCC	CTGACAACAT	CAACTGGTCC	2400
	CAGTTTATTC	CTGAGCTACA	GTAGAGCCCT	GCCCTTGCCC	CTGTGCTCAA	GCTGTCCACC	2460
	ATCCCGGGCA	CTCCAAGGCT	CAGTGCACCC	CAAGCCTCTG	AGTGAGGACA	GCAGGCAGGG	2520
70	ACTGTTCTGC	TCCTCATAGC	TCCCTGCTGC	CTGATTATGC	AAAAGTAGCA	GTCACACCCCT	2580
	AGCCACTGCT	GGGACCTTGT	GTTCCCCAAG	AGTATCTGAT	TCCTCTGCTG	TCCCTGCCAG	2640
	GAGCTGAAGG	GTGGGAACAA	CAAAGCAAT	GGTGAAAGA	GATTAGGAAC	CCCCCAGCCT	2700
	GTTTCCATTC	TCTGCCAGC	AGTCTCTTAC	CTTCCCTGAT	CTTTGCAGGG	TGGTCCGTGT	2760
	AAATAGTATA	AATTCTCAA	ATTATCCTCT	AATTATAAAT	GTAAGCTTAT	TTCCTTAGAT	2820
75	CATTATCCAG	AGACTGCCAG	AAGGTGGGTA	GGATGACCTG	GGGTTTCAAT	TGACTTCTGT	2880
	TCCTTGCTTT	TAGTTTGTAT	AGAAGGGAAG	ACCTGCAGTG	CACGGTTTCT	TCCAGGCTGA	2940
	GGTACCTGGA	TCTTGGGTTC	TTCACCTGAG	GGACCCAGAC	AAGTGGATCT	GCTTGCCAGA	3000
	GTCTTTTGTG	CCCTTCCCTG	CCACTTCCC	GTGTTTCCAA	GTCAGCTTTC	CTGCAAGAAG	3060
	AAATCCTGGT	TAAAAAAGTC	TTTTGTATTG	GGTCAGGAGT	TGAATTTGGG	GTGGGAGGAT	3120
80	GGATGCAACT	GAAGCAGAGT	GTGGGTGCCC	AGATGTGCGC	TATTAGATGT	TTCTCTGATA	3180
	ATGTCCCCAA	TCATACCAGG	GAGACTGGCA	TTGACGAGAA	CTCAGGTGGA	GGCTTGAGAA	3240
	GGCCGAAAGG	GCCCCTGACC	TGCCTGGCTT	CCTTAGCTTG	CCCCTCAGCT	TTGCAAAGAG	3300
	CCACCCTAGG	CCCCAGCTGA	CCGCATGGGT	GTGAGCCAGC	TTGAGAACAC	TAACACTCA	3360
85	ATAAAAGCGA	AGGTGGAAAA	AAAAAAAAAA	AAAAAA			

Seq ID NO: 71 Protein sequence:  
Protein Accession #: AAH06529.1



	1	11	21	31	41	51	
5	MKTSRRRLPLI	LKRRRLPLPV	QNAPSETSEE	EPKRSPAQQE	SNQAEASKEV	AESNSCKFPA	60
	GIKIINHPTM	PNTQVVAIPN	NANIHSIITA	LTAKGKESGS	SGPNKFILIS	CGGAPTQPPG	120
	LRPQTQTSYD	AKRTEVTLET	LGPKEAARDV	NLPRPPGALC	EQKRETCADG	EAAGCTINNS	180
	LSNIQWLRKM	SSDGLGSRSI	KQEMEKEKENC	HLEQRQVKVE	EPSRPSASWQ	NSVSRPPYS	240
	YMAMIQFAIN	STERKRMTLK	DIYTWIEDHF	PYFKHIAKPG	WKNSIRHNL	LHDMFVRETS	300
10	ANGKVSFWTI	HPSANRYLTL	DQVFKQKRP	NPELRRNMTI	KTELPLGARR	KMKPLLPRVS	360
	SYLVPIQFPV	NQSLVLQPSV	KVPLPLAASL	MSSELARHSK	RVRIAPKVLL	AEEGIAPLSS	420
	AGPGKEEKLL	FGEGFSPLLP	VQTIKEEEIQ	PGEEMPHLAR	PIKVESPPLE	EWSPAPSPFK	480
	EESHSWEDS	SQSPTPRPKK	SYSGLRSPTR	CVSEMLVIQH	RERRERSRSR	RKQHLLPPCV	540
	DEPELLFSEG	PSTSRWAAEL	PFADSSDPA	SQLSYSQEVG	GPFKTPIKET	LPISSTPSKS	600
15	VLPRTPEWR	LTPPAKVGL	DFSPVQTPQG	ASDPLDPLG	LMDLSTTPLQ	SAPPLESPQR	660
	LLSSEPLDLI	SVDPGNSSPS	DIDVPKPGSP	EPQVSGLAAN	RSLTEGLVLD	TMNDSLSKIL	720
	LDISFPGLE	DPLGPDNINW	SQFIPELQ				

20 Seq ID NO: 72 DNA sequence  
Nucleic Acid Accession #: U74612.1  
Coding sequence: 178-2583

	1	11	21	31	41	51	
25	GGCACGAGGG	GGACCCGGCC	GGTCCGGCGC	GAGCCCCCGT	CCGGGGCCCT	GGCTCGGCCC	60
	CCAGGTTGGA	GGAGCCCGGA	GCCCCGCTTC	GGAGCTACGG	CCTAACGGCG	GCGGCGACTG	120
	CAGTCTGGAG	GGTCCACACT	TGTGATTCTC	AATGGAGAGT	GAAAACGCAG	ATTCATAATG	180
	AAAAC TAGCC	CCCGTCGGCC	ACTGATTCTC	AAAAGACGGA	GGCTGCCCCC	TCCTGTTCAA	240
	AATGCCCCAA	GTGAAACATC	AGAGGAGGAA	CCTAAGAGAT	CCCCTGCCCA	ACAGGAGTCT	300
30	AATCAAGCAG	AGGCCTCCAA	GGAAGTGGCA	GAGTCCAACT	CTTGCAAGTT	TCCAGCTGGG	360
	ATCAAGATTA	TTAACCACCC	CACCATGCCC	AACACGCAAG	TAGTGGCCAT	CCCCAACAA	420
	GCTAATATTC	ACAGCATCAT	CACAGCACTG	ACTGCCAAGG	GAAAAGAGAG	TGGCAGTAGT	480
	GGGCCCCAACA	AATTCATCCT	CATCAGCTGT	GGGGGAGCCC	CAACTCAGCC	TCCAGGACTC	540
	CGGCCCTCAA	CCCAAACCCAG	CTATGATGCC	AAAAGGACAG	AAGTGACCCT	GGAGACCTTG	600
35	GGACCAAAAC	CTGCAGCTAG	GGATGTGAAT	CTTCCTAGAC	CACCTGGAGC	CCTTTGCGAG	660
	CAGAAACGGG	AGACCTGTGC	AGATGGTGAG	GCAGCAGGCT	GCACTATCAA	CAATAGCCTA	720
	TCCAACATCC	AGTGGCTTCG	AAAGATGAGT	TCTGATGGAC	TGGGCTCCCG	CAGCATCAAG	780
	CAAGAGATGG	AGGAAAAGGA	GAATTGTGAC	CTGGAGCAGC	GACAGGTTAA	GGTTGAGGAG	840
	CCTTCGAGAC	CATCAGCGTC	CTGGCAGAAC	TCTGTGTCTG	AGCGGCCACC	CTACTCTTAC	900
40	ATGGCCATGA	TACAATTCGC	CATCAACAGC	ACTGAGAGGA	AGCGCATGAC	TTTGAAAGAC	960
	ATCTATACGT	GGATTGAGGA	CCACTTTCCC	TACTTTAAGC	ACATTGCCAA	GCCAGGCTGG	1020
	AAGAACTCCA	TCCGCCACAA	CCTTTCCCTG	CACGACATGT	TTGTCCGGGA	GACGTCTGCC	1080
	AATGGCAAGG	TCTCCTTCTG	GACCATTAC	CCCAGTGCCA	ACCGCTACTT	GACATTGGAC	1140
	CAGGTGTTTA	AGCCACTGGA	CCCAGGGTCT	CCACAATTGC	CCGAGCACTT	GGAATCACAG	1200
45	CAGAAACGAC	CGAATCCAGA	GCTCCGCGCG	AACATGACCA	TCAAACCGA	ACTCCCCCTG	1260
	GGCGCACGGC	GGAAAGTGAA	GCCAGTGCTA	CCACGGGTCA	GCTCATACCT	GGTACCTATC	1320
	CAGTTCCCGG	TGAACAGTGC	ACTGGTGTGG	CAGCCCTCGG	TGAAGGTGCC	ATTGCCCTTG	1380
	GCGGCTTCCC	TCATGAGCTC	AGAGCTTGCC	CGCCATAGCA	AGCGAGTCCG	CATTGCCCCC	1440
	AAGGTTTTTG	GGGAACAGGT	GGTGTGTTGG	TACATGAGTA	AGTTCTTTAG	TGGCGATCTG	1500
50	CGAGATTTTG	GTACACCCAT	CACCAGCTTG	TTTAATTTTA	TCTTTCTTTG	TTTATCAGTG	1560
	CTGCTAGCTG	AGGAGGGGAT	AGCTCCTCTT	TCTTCTGCAG	GACCAGGGAA	AGAGGAGAAA	1620
	CTCCTGTTTG	GAGAAGGGTT	TTCTCCTTTG	CTTCCAGTTC	AGACTATCAA	GGAGGAAGAA	1680
	ATCCAGCCTG	GGGAGGAAAT	GCCACACTTA	GCGAGACCCA	TCAAAGTGGA	GAGCCCTCCC	1740
	TTGGAAGAGT	GGCCCTCCCC	GGCCCATCT	TTCAAAGAGG	AATCATCTCA	CTCCTGGGAG	1800
55	GATTCGTCCC	AATCTCCAC	CCCAAGACCC	AAGAAGTCCT	ACAGTGGGCT	TAGGTCCCCA	1860
	ACCCGCTGTG	TCTCGGAAAT	GCTTGTGATT	CAACACAGGG	AGAGGAGGGA	GAGGAGCCGG	1920
	TCTCGGAGGA	AACAGCATCT	ACTGCCTCCC	TGTGTGGATG	AGCCGGAGCT	GCTCTTCTCA	1980
	GAGGGGCCCA	GTACTTCCCG	CTGGGCCGCA	GAGTCCCGT	TCCCAGCAGA	CTCCTCTGAC	2040
	CCTGCCCTCCC	AGCTCAGCTA	CTCCAGGAA	TGGCGAGGAC	CTTTTAAGAC	ACCCATTAA	2100
60	GAAACGCTGC	CCATCTCCTC	CACCCCGAGC	AAATCTGTCC	TCCCCAGAAC	CCCTGAATCC	2160
	TGGAGGCTCA	CGCCCCCAGC	CAAAGTAGGG	GGACTGGATT	TCAGCCCAGT	ACAAACCTCC	2220
	CAGGGTGCC	CTGACCCCTT	GCCTGACCCC	CTGGGGCTGA	TGGATCTCAG	CACCACTCCC	2280
	TTGCAAGTG	CTCCCCCCTT	TGAATCACCG	CAAAGGCTCC	TCAGTTTACA	ACCCTTAGAC	2340
	CTCATCTCCG	TCCCTTTTGG	CAACTCTTCT	CCCTCAGATA	TAGACGTCCC	CAAGCCAGGC	2400
65	TCCCGGAGC	CACAGGTTTC	TGGCCTTGCA	GCCATTCGTT	CTCTGACAGA	AGGCCTGGTC	2460
	CTGGACACAA	TGAATGACAG	CCTCAGCAAG	ATCTGTCTGG	ACATCAGCTT	TCCTGGCCTG	2520
	GACGAGGACC	CACTGGGCCC	TGACAACATC	AACTGGTCCC	AGTTTATTCC	TGAGCTACAG	2580
	TAGAGCCCTG	CCCTTGCCCC	TGTGCTCAAG	CTGTCCACCA	TCCCGGGCAC	TCCAAGGCTC	2640
	AGTGACCCCC	AAGCCTCTGA	GTGAGGACAG	CAGGCAAGGA	CTGTTCTGCT	CCTCATAGCT	2700
70	CCCTGCTGCC	TGATTATGCA	AAAGTAGCAG	TCACACCCTA	GCCACTGCTG	GGACCTTGTG	2760
	TTCCCCAAGA	GTATCTGATT	CCTCTGCTGT	CCCTGCCAGG	AGCTGAAGGG	TGGGAACAAC	2820
	AAAGGCAATG	GTGAAAAGAG	ATTAGGAACC	CCCCAGCCTG	TTTCCATTCT	CTGCCAGCA	2880
	GTCTCTTACC	TTCCCTGATC	TTTGAGGGT	GGTCCGTGTA	AATAGTATAA	ATTCTCCAAA	2940
	TTATCCTCTA	ATTATAAATG	TAAGCTTATT	TCCTTAGATC	ATTATCCAGA	GACTGCCAGA	3000
75	AGGTGGGTAG	GATGACCTGG	GGTTTCAATT	GACTTCTGTT	CCTTGCTTTT	AGTTTGTGATA	3060
	GAAGGGAAGA	CCTGCAGTGC	ACGGTTTCTT	CCAGGCTGAG	GTACCTGGAT	CTTGGGTTCT	3120
	TCATGACAGG	GACCCAGACA	AGTGGATCTG	CTTGCCAGAG	TCCTTTTTGC	CCCTCCCTGC	3180
	CACCTCCCCG	TGTTTCCAAG	TCAGGATTTCC	TGCAAGAAGA	AATCCTGGTT	AAAAAAGTCT	3240
	TTTGTATTGG	GTCAGGAGTT	GAATTGTTGGG	TGGGAGGATG	GATGCAACTG	AAGCAGAGTG	3300
80	TGGGTGCCCA	GATGTGCGCT	ATTAGATGTT	TCTCTGATAA	TGTCCCCAAT	CATACCAGGG	3360
	AGACTGGCAT	TGACGAGAAC	TCAGGTGGAG	GCCTTGAGAAG	GCCGAAAGGG	CCCCTGACCT	3420
	GCCTGGCTTC	CTTAGCTTGC	CCCTCAGCTT	TGCAAGAGAGC	CACCCTAGGC	CCCAGCTGAC	3480
	CGCATGGGTG	TGAGCCAGCT	TGAGAACACT	AACTACTCAA	TAAAAGCGAA	GGTGGACAAA	3540
85	AAAAAAAAAA	AAAAA					

Seq ID NO: 73 Protein sequence:  
Protein Accession #: AAC51128.1



	1	11	21	31	41	51	
5	MKTSRRPLI	LKRRRLPLPV	QNAPSETSEE	EPKRSPAQQE	SNQAEASKEV	AESNSCKFPA	60
	GIKIINHPTM	PNTQVVAIPN	NANIHSIITA	LTAKGKESGS	SGPNKFILIS	CGGAPTQPPG	120
	LRPQTQTSYD	AKRTEVTLET	LGPKPAARDV	NLPRPPGALC	EOKRETCADG	EAAGCTINNS	180
	LSNIQWLRKM	SSDGLGSRSI	KQEMEKEKENC	HLEQRQVKVE	EPSRPSASWQ	NSVSRPPYS	240
	YMAMIQFAIN	STERKRMTLK	DIYTWIEDHF	PYFKHIAKPG	WKNSIRHNLS	LHDMFVRETS	300
10	ANGKVSFWTI	HPSANRYLTL	DQVFKPLDPG	SPQLPEHLES	QOKRPNPELR	RNMTIKTELP	360
	LGARRKMKPL	LPRVSSYLVP	IQFPVNQSLV	LQPSVKVPLP	LAASLMSEL	ARHSKRVRIA	420
	PKVFGEQVVF	GYMSKFFSGD	LRDFGTPITS	LFNFIFLCLS	VLLAEEGIAP	LSSAGPGKEE	480
	KLLFGEFSP	LLPVQTIKEE	EIQPGEEMPH	LARPIKVESP	PLEEWSPAP	SFKEESSHSW	540
	EDSSQSPTPR	PKKSYSGLRS	PTRCVSEMLV	IQHRERRERS	RSRRKQHLLP	PCVDEPELLF	600
15	SEGPSTSRWA	AELPFPADSS	DPASQLSYSQ	EVGGPFKTP	KETLPISSTP	SKSVLPRTPE	660
	SWRLTPPAKV	GGLDFSPVQT	SQGASDPLPD	PLGLMDLSTT	PLQSAPPLES	PQRLLSSEPL	720
	DLISVPFGNS	SPSDIDVPKP	GSPEPQVSGL	AANRSLTEGL	VLDTMNDSLS	KILLDISFPG	780
	LDEDPLGPDN	INWSQFIPEL	Q				

Seq ID NO: 74 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 111-416

	1	11	21	31	41	51	
25	GGGAAGAGCC	AGGCTGAGCC	TTATAAAGGA	CTGCTCTTTG	TCCAAACACA	CACATCTCAC	60
	TCATCCTTCT	ACTCGTGACG	CTTCCCAGCT	CTGGCTTTTT	GAAAGCAAAG	ATGAGCAACA	120
	CTCAAGCTGA	GAGGTCCATA	ATAGGCATGA	TCGACATGTT	TCACAAATAC	ACCAGACGTG	180
	ATGACAAGAT	TGAGAAGCCA	AGCCTGCTGA	CGATGATGAA	GGAGAACTTC	CCCAACTTCC	240
30	TTAGTGCCTG	TGACAAAAAG	GGCACAATT	ACCTCGCCGA	TGTCTTTGAG	AAAAAGGACA	300
	AGAATGAGGA	TAAGAAGATT	GATTTTTCTG	AGTTTCTGTC	CTTGCTGGGA	GACATAGCCA	360
	CAGACTACCA	CAAGCAGAGC	CATGGAGCAG	CGCCCTGTTC	CGGGGGCAGC	CAGTGACCCA	420
	GCCCCACCAA	TGGGCCTCCA	GAGACCCCAG	GAACAATAAA	ATGTCTTCTC	CCACCAGA	

Seq ID NO: 75 Protein sequence:  
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
40	MSNTQAERSI	IGMIDMFHKY	TRRDDKIEKP	SLLTMMKENF	PNFLSACDKK	GTNYLADVFE	60
	KKDKNEDKKI	DFSEFLSLLG	DIATDYHKQS	HGAAPCSGGS	Q		

Seq ID NO: 76 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 111-416

	1	11	21	31	41	51	
50	GGGAAGAGCC	AGGCTGAGCC	TTATAAAGGA	CTGCTCTTTG	TCCAAACACA	CACATCTCAC	60
	TCATCCTTCT	ACTCGTGACA	CTTCCCAGTT	CTGGCTTTTT	GAAAGCAAAG	ATGAGCAACA	120
	CTCAAGCTGA	GAGGTCCATA	ATAGGCATGA	TCGACATGTT	TCACAAATAC	ACCGGACGTG	180
	ATGGCAAGAT	TGAGAAGCCA	AGCCTGCTGA	CGATGATGAA	GGAGAACTTC	CCCAATTTC	240
55	TCAGTGCCTG	TGACAAAAAG	GGCACAATT	ACCTCGCCAC	TGTCTTTGAG	AAAAAGGACA	300
	AGAATGAGGA	TAAGAAGATT	GATTTTTCTG	AGTTTCTGTC	CTTGCTGGGA	GACATAGCCG	360
	CAGACTACCA	CAAGCAGAGC	CATGGAGCGG	CGCCCTGTTC	TGGGGGAAGC	CAGTGATCCA	420
	GCCCCACCAA	GGGGCCTCCA	GAGACCCCAG	GAACAATAAG	TGTCTCTCTC	CACCAGA	

Seq ID NO: 77 Protein sequence:  
Protein Accession #: XP\_048124.1

	1	11	21	31	41	51	
60	MSNTQAERSI	IGMIDMFHKY	TGRDGKIEKP	SLLTMMKENF	PNFLSACDKK	GIHYLATVFE	60
65	KKDKNEDKKI	DFSEFLSLLG	DIAADYHKQS	HGAAPCSGGS	Q		

Seq ID NO: 78 DNA sequence  
Nucleic Acid Accession #: Z73678.1  
Coding sequence: 253-2433

	1	11	21	31	41	51	
75	GGGGTGGTGC	AGGGCAGGGG	TGGTATATCC	TGTCTGACGG	AGGGCGGGCC	TCGCCAGTGC	60
	CAGAGAGGGA	CGAACCAGGG	TGGAAGCGCC	AGGAGCAGCT	GCAGGGAGCC	CTCACGCGGA	120
	CCTCGCACTC	TATGGCCGTA	GGGAGCCGCT	GAGAGCGAGA	AGAGCACGCT	CCTGCCCCGC	180
	CGCTGCACCG	CACCTCGCCT	CGCCTCTCTG	CTCTCCTAGG	CCCCGGCCGC	GCGCCACCCG	240
	CCTCCCGCCA	CCATGAACCA	CTCGCCGCTC	AAGACCGCCT	TGGCGTACGA	ATGCTTCCAG	300
80	GACCAAGACA	ACTTCCAGTT	GGCTTTCGCG	TCCGACCAAA	AGATGAAAAC	AGGCACGTCT	360
	GGCAGGCAGC	GCCTGCAGGA	GCAGGTGATG	ATGACCGTCA	AGCGGCAGAA	GTCCAAGTCT	420
	TCCCAGTCGT	CCACCCTGAG	CCACTCCAAT	CGAGGTTCCT	TGTATGATGG	CTTGGCTGAC	480
	AATTACAAC	ATGGGACCAC	CAGCAGGAGC	AGCTACTACT	CCAAGTTCCA	GGCAGGGAAT	540
	GGCTCATGGG	GATATCCGAT	CTACAATGGA	ACCTCAAGC	GGGAGCCTGA	CAACAGGCGC	600
	TTCAGCTCCT	ACAGCCAGAT	GGAGAATCTG	AGCCGGCACT	ACCCCGGGG	CAGCTGTAAC	660
85	ACCACCGGCG	CAGGCAGCGA	CATCTGCTTC	ATGCAGAAAA	TCAAGGCGAG	CCGCAGTGAG	720
	CCCAGCTCT	ACTGTGACCC	ACGGGGCACC	CTGCGCAAGG	GCACGCTGGG	CAGCAAGGGC	780
	CAGAAGACCA	CACAGAACCG	CTACAGCTTT	TACAGCACCT	GCAGTGGTCA	GAAGGCCATA	840
	AAGAAGTGCC	CTGTGCGCCC	GCCCTCTTGT	GCCTCCAAGC	AGGACCCTGT	GTATATCCCG	900



	CCCATCTCCT	GCAACAAGGA	CCTGTCCTTT	GGCCACTCTA	GGGCCAGCTC	CAAGATCTGC	960
	AGTGAGGACA	TCGAGTGCAG	TGGGCTGACC	ATCCCCAAGG	CTGTGCAGTA	CCTGAGCTCC	1020
	CAGGATGAGA	AGTACCAGGC	CATTGGGGCC	TATTACATCC	AGCATACCTG	CTTCCAGGAT	1080
5	GAATCTGCCA	AGCAACAGGT	CTATCAGCTG	GGAGGCATCT	GCAAGCTGGT	GGACCTCCTC	1140
	CGCAGCCCCA	ACCAGAACGT	CCAGCAGGCC	GCGGCAGGGG	CCCTGCGCAA	CCTGGTGTTC	1200
	AGGAGCACCA	CCAACAAGCT	GGAGACCCGG	AGGCAGAAATG	GGATCCGCGA	GGCAGTCAGC	1260
	CTCCTGAGGA	GAACCGGGAA	CGCCGAGATC	CAGAAGCAGC	TGACTGGGCT	GCTCTGGAAC	1320
	CTGTCTTCCA	CTGACGAGCT	GAAGGAGGAA	CTCATTGCCG	ACGCCCTGCC	TGTTCTGGCC	1380
10	GACCGCGTCA	TCATTCCCTT	CTCTGGCTGG	TGCGATGGCA	ATAGCAACAT	GTCCCGGGAA	1440
	GTGGTGGACC	CTGAGGTCTT	CTTCAATGCC	ACAGGCTGCT	TGAGGAACCT	GAGCTCGGCC	1500
	GATGCAGGCC	GCCAGACCAT	GCGTAACTAC	TCAGGGCTCA	TTGATTCCCT	CATGGCCTAT	1560
	GTCCAGAACT	GTGTAGCGGC	CAGCCGCTGT	GACGACAAGT	CTGTGGAAAA	CTGCATGTGT	1620
	GTTCTGCACA	ACCTCTCCTA	CCGCCTGGAC	GCCGAGGTGC	CCACCCGCTA	CCGCCAGCTG	1680
	GAGTATAACG	CCCCGAACGC	CTACACCGAG	AAGTCCCTCA	CTGGCTGCTT	CAGCAACAAG	1740
15	AGCGACAAGA	TGATGAACAA	CAACTATGAC	TGCCCCCTGC	CTGAGGAAGA	GACCAACCCC	1800
	AAGGGCAGCG	GCTGGTTGTA	CCATTTCAGAT	GCCATCCGCA	CCTACCTGAA	CCTCATGGGC	1860
	AAGAGCAAGA	AAGATGCTAC	CCTGGAGGCC	TGTGCTGGTG	CCCTGCAGAA	CCTGACAGCC	1920
	AGCAAGGGGC	TGATGTCCAG	TGGCATGAGC	CAGTTGATTG	GGCTGAAGGA	AAAGGGCCTG	1980
20	CCACAAATTG	CCCGCCTCCT	GCAATCTGGC	AACTCTGATG	TGGTGCGGTC	CGGAGCCTCC	2040
	CTCCTGAGCA	ACATGTCCCG	CCACCCTCTG	CTGCACAGAG	TGATGGGGAA	CCAGGTGTTC	2100
	CCGGAGGTGA	CCAGGCTCCT	CACCAGCCAC	ACTGGCAATA	CCAGCAACTC	CGAAGACATC	2160
	TTGTCCCTCG	CCTGCTACAC	TGTGAGGAAC	CTGATGGCCT	CGCAGCCACA	ACTGGCCAAG	2220
	CAGTACTTCT	CCAGCAGCAT	GCTCAACAAC	ATCATCAACC	TGTGCCGAAG	CAGTGCCTCA	2280
25	CCCAAGGCCG	CAGAAGCTGC	CCGGCTTCTC	CTGTCTGACA	TGTGGTCCAG	CAAGGAAGCTG	2340
	CAGGGTGTCC	TCAGACAGCA	AGGTTTCGAT	AGGAACATGC	TGGGAACCTT	AGCTGGGGCC	2400
	AACAGCCTCA	GGAACCTTAC	CTCCCGATTG	TAAGAAGAGA	CTGTCCAAGC	AAGTTAGGCT	2460
	TGCAGGAAGA	TATGACCCAG	CTGAGAAGCC	CTCAGGCCTC	GCTGGATGGG	GTTTTCTGTC	2520
	CATCCTGTGC	AGTATTTGGG	AAAGTTTACA	AGAAACTGAG	AAGAAACCTA	AAAAGTGTGG	2580
30	ATAGTGGAAA	GATTTTGTAG	TTTTTTTTTT	CCTTGGGGAA	ACTGGCAGGC	AATGGGGGTT	2640
	AGGGAGGTTG	GGGCGGGGGG	GGCTTTCTTG	AGTTAAAGGG	GCTTATATGT	GATGTCAATA	2700
	TTTCTTCCCT	TGAGAAATGG	TATATATATG	TGTCTAATGT	AAGTGTGTGC	ATGCATGTGC	2760
	GCGTGCATGT	GTGTGTGTGT	GAGTGTCTTA	AAGCATAACC	ACAAACTGCA	AAAAGCTAGG	2820
	TAAGCTATTT	TGTTGCAGCT	CATAAGGTGG	TGAAAAGGAC	TCTCCTGTGT	TTCTTACTCA	2880
35	TAGGCAAGGA	CAACATGTGC	TTTTTGGTGA	GCTGCTCATA	ATTCCTGAAA	TGTGTGGTGC	2940
	CAGGGCAAGG	GGGCCATCAC	TGCAGTCAGG	CCCTCAGAGG	AGTCTGCAG	GCTTCTTACC	3000
	AGTGGTCTCC	AAGGGTGCAG	GAGTAACTGG	GGCTGGGCCA	GCCTCCCCCC	TTACAAGGCT	3060
	GCTTTCCACG	AAGGGAGGTC	TGGTGTATCT	CATGGGAGAA	TCTGGGGTGT	CTGTAGTGTG	3120
	ACCCCTCCAG	CAGCGCCACA	AGGACTGAGG	TTGGGTAGGT	GTGAGGTTCC	AGAGGACAGC	3180
40	AGGACACTCT	CGCATACTTT	GCCAAATGAG	GCCTGCTCAG	AGGAGTAGGA	GCTGAAAGAT	3240
	GGTGCCTTCC	ACCCTCTTGG	GCTGTGTGCC	CTCAGAGCA	GGCTCAGCCT	GCAAAGGCC	3300
	TGCATTTCAG	GGTCTTGTAA	TCTACTTGTT	GCAGGAGAAA	GAAGGTAAAA	AATGATTTTT	3360
	TTAAGAAAAG	CTATTTTATT	GCAGCTCTTT	CCCAAGAGCT	GTTCTGGGAA	TGGCTGGTCT	3420
	TCATATTCCC	AGTGGAGAGG	GGAACAAGTG	GGGCTGGGCA	TATACCTATT	CCGGCTTCTA	3480
45	GTGGGATGGA	GTTGGGTAT	AGAAATTAAC	CAGGAAGATG	TTTCCACCAA	GCCTGCTGTG	3540
	AGTCAATTGA	GGGAGTGTTT	GGGTCCCAGG	AGACTTGGAC	GGGGGGAGTT	TGGGTAGACT	3600
	AGGAAAGGAA	AGTGCCATAT	CAGGGTACCG	GTACTCGGCA	GCTCACATCT	CAGCCAGGGG	3660
	CCATGCCCCA	CTTCCCTGTA	CCCCAGCTGT	CTTGTCTCCA	CTCTGTGAAA	CCCACAGGGG	3720
	ATGTGATAAA	CAGGGCTATT	AGGGGTATCA	GCCACGTCGA	GCCCCCAGAC	TCTGTGCACT	3780
50	TCAGACCAGC	AGCAGCAGGA	GGGCTCCCGA	GGGCCCTTATG	AGAAAACCTG	TGTGGACATC	3840
	CCTTGGTGTG	CACATAAGACA	GAGCAGAGCC	CAGCGCTCCC	AAGCCTTCCT	CCTTCCAGCT	3900
	TCTACCTCCA	TGCTAGCATT	GCTGGTGTTA	GAGAGGAATT	AACTTCTTGG	TCTGTGCCCT	3960
	TCTCTAGAAG	AATATAAGAT	GCTCCTCCTC	CTCACCCCTT	CTCAGCCTCC	TCCCAAGTCT	4020
	TCTCTTCTG	CACCACCCCC	GAGTCCAAAC	CCACCTCTTG	CCCCCAGATT	CAGGCTGGAA	4080
55	AACACTGATG	TGGACTCAGT	ATGACAACCTG	AGATGGGGGA	AGCCAGACAT	GTGAGGACGC	4140
	TGTCTCCGA	GAGGTGTCCC	CGGCTGTTAG	CCAGCTGTGC	TGTGGTGCTG	TGGGTCTGTC	4200
	ATACCCTCCC	TTGCTTCTGT	TCACACTGGG	AGGCCCACTC	CTGGCTCACC	TCTCCCTCTC	4260
	AGGGACCCAC	GTGGGAGCCT	GGATCCCTGG	ACTGTCTCTG	GCATAGGTTT	CAGGGGCCTC	4320
	CTTTGTTGTC	ATCAGAACCC	AGAGGAATTG	TTCTCCTAAA	AAATACGTAT	GGCATAACCA	4380
60	TCTGTGCGGG	GCAGTGTCTT	AAGCACTTAG	ACTACATCAG	GGAAGAACAC	AGACCACATC	4440
	CCCGTCCCTCA	TGCGGCTTAT	GTTTTCTGGA	GGAAAGTGGG	GACACAAGTC	CTTGGCTTTA	4500
	GGGCTCCCCC	GGCTGGGGGC	TGTGCAGTCC	GGTCAGGGCG	GGAGGGGAAA	TGCACCGCTG	4560
	CATGTGAACC	TTACCAGCCC	AGGCGGATGC	CCCTTCCCTT	TAGCACTACC	CTGGCCTCCT	4620
	GCATCCCTCT	GCCTCATGTT	CCTCCACACT	TCAAAGAATG	AAGAGCCCCA	TGGGCCCAGC	4680
65	CCCTGCCCTG	GGAACCAGGC	AGCCTTCCAG	ACCTCAGGGG	CTGAGGCAGA	CTATTAGGGC	4740
	AGGGCTGACT	TTGGTGACAC	TGCCCCATTCC	CTCTCAGGCC	AGCTCAGGTC	ACCCGGGCCT	4800
	CTGACCCAGG	CTGTCACTT	TGAGAGGGGC	AAAAGTGAAG	GGGGCTTTTC	CTAGAGAAAG	4860
	AGAACAAGGA	GCTTGCCAGG	CTTCATGTAG	CCGACACACG	TCTCAGGATT	TTAAGTCCAC	4920
	ATTGGCCTCA	CACAGCCTA	GGCCAATGCC	CAAAATAAGG	AGTTCCAATT	TGGGGCCAAA	4980
70	TGAGGAAGGA	CACAGACTCT	GCCCTGGGAT	CTCCTGTGCT	AGCGGCCAAT	GACAAATCCA	5040
	GTCATTGGCC	ACCAGCCACC	PCTGCAGTGG	GGACCACACT	AGCAGCCCTG	ACTCCACACT	5100
	CCTCCTGGGG	ACCAAGAGG	CAGTGTGTGT	GTCTGCGTGT	CCACCTTGA	ATCTGGCTGA	5160
	ACTGGCTGGG	AGGACCAAGA	CTGCGGCTGG	GTGGGGCAGG	GAAGGGGAAGC	CGGGGGCTGC	5220
	TGTGAGGGAT	CTTGAGGCTT	CCCTGTAGCC	CACCTTCCCC	TTGCTTCATG	TTTGTAGAGG	5280
75	AACCTTGTGC	CGGCCAGGCC	CAGTTTCTTT	GTGTGATACA	CTAATGTATT	TGCTTTTTTT	5340
	GGAAATAGAG	AAAATCAATA	AATTGCTAGT	GTTTCTTTGA	AAAAAATA		

Seq ID NO: 79 Protein sequence:  
Protein Accession #: CAA98022.1

80	1	11	21	31	41	51	
	MNHSPLKTAL	AYECFQDQDN	STLALPSDQK	MKTGTSGRQR	VQEQVMMTVK	RQKSKSSQSS	60
	TLSHSNRGS	YDGLADNYNY	GTTSRSSYY	KFQAGNGSWG	YPIYNGTLKR	EPDNRFRSSY	120
	SQMNWSRHY	PRGSCNTTGA	GSDICFMQKI	KASRSEPDLY	CDPRGTLRKG	TLGSKGQKTT	180
85	QNRYSFYSTC	SGQKAIKKCP	VRPPSCASKQ	DPVYIPPIISC	NKDLSFGHSR	ASSKICSEDI	240
	ECSGLTIPKA	VQYLSSQDEK	YQAIGAYYIQ	HTCFQDESAK	QQVYQLGGIC	KLVDLLRSPN	300
	QNVQQAAGA	LRNLVFRSTT	NKLETRRQNG	IREAVSLLRR	TGNAEIQKQL	TGLLWNLSST	360



DELKEELIAD ALPVLADRV IIPFSGWCDGN SNMSREVVDP EVFFNATGCL RNLSSADAGR 420  
QTMRYNSGLI DSLMAYVQNC VAASRCDDKS VENCMLVHNS LSYRLDAEVP TRYRQLEYNA 480  
RNAYTEKSST GCFSNKSDDK MNNNYDCPLP EETNPKGSG WLYHSDAIR YLNLMGKSKK 540  
5 DATLEACAGA LQNLTASKGL MSSGMSQLIG LKEKGLPQIA RLLQSGNSDV VRSGASLLSN 600  
MSRHPLLRV MGNQVPFVET RLLTSHGTNT SNEEDILSSA CYTVRNLMAS QPQLAKQYFS 660  
SSMLNNIINL CRSSASPKAA EAARLLLSDM WSSKELQGV L RQQGFDRNML GTLAGANSLR 720  
NFTSRF

Seq ID NO: 80 DNA sequence  
Nucleic Acid Accession #: NM\_006516.1  
Coding sequence: 180-1658

1 11 21 31 41 51  
| | | | | |  
15 TAGTCGCGGG TCCCCGAGTG AGCACGCCAG GGAGCAGGAG ACCAAACGAC GGGGGTCGGA 60  
GTCAGAGTCG CAGTGGGAGT CCCCAGGACG GAGCAGGAGC CTGAGCGGGA GAGCGCCGCT 120  
CGCAGCGCCG TCGCCACCCG CGTACCCGGC GCAGCCAGAG CCACCAGCGC AGCGCTGCCA 180  
20 TGGAGCCCAG CAGCAAGAAG CTGACGGGTC GCCTCATGCT GGCTGTGGGA GGAGCAGTGC 240  
TTGGCTCCCT GCAGTTTGGC TACAACACTG GAGTCATCAA TGCCCCCAG AAGGTGATCG 300  
AGGAGTTCTA CAACCAGACA TGGGTCCACC GAGTCATCCTG CCCACCACGC 360  
TCACCACGCT CTGGTCCCTC TCAGTGGCCA TCTTTTCTGT TGGGGGCATG ATTGGCTCCT 420  
TCTCTGTGGG CTTTTTCGTT AACCGCTTTG GCCGGCGGAA TTCAATGCTG ATGATGAACC 480  
TGCTGGCCTT CGTGTCCGCC GTGCTCATGG GCTTCTCGAA ACTGGGCAAG TCCTTTGAGA 540  
25 TGCTGATCCT GGGCCGCTTC ATCATCGGTG TGTACTGCGG CCTGACCACA GGCTTCGTGC 600  
CCATGTATGT GGGTGAAGTG TCACCCACAG CCTTTCGTGG GGCCCTGGGC ACCCTGCACC 660  
AGCTGGGCAT CGTCGTCGGC ATCCTCATCG CCCAGGTGTT CGGCCTGGAC TCCATCATGG 720  
GCAACAAGGA CCTGTGGCCC CTGCTGCTGA GCATCATCTT CATCCCGGCC CTGCTGCAGT 780  
GCATCGTGCT GCCCTTCTGC CCCGAGAGTC CCCGCTTCCT GCTCATCAAC CGCAACGAGG 840  
30 AGAACCAGGC CAAGAGTGTG CTAAAGAAGC TGCGCGGGAC AGCTGACGTG ACCCATGACC 900  
TGCAAGAGAT GAAGGAAGAG AGTCGGCAGA TGATGCGGGA GAAGAAGGTC ACCATCCTGG 960  
AGCTGTTCCG CTCCCCCGCC TACCGCCAGC CCATCCTCAT CGCTGTGGTG CTGCAGCTGT 1020  
CCCAGCAGCT GTCTGGCATC AACGCTGTCT TCTATTACTC CACGAGCATC TTCGAGAAGG 1080  
CGGGGGTGCA GCAGCCTGTG TATGCCACCA TTGGTCCCGG TATCGTCAAC ACGGCCTTCA 1140  
35 CTGTCGTGTC GCTGTTTGTG GTGGAGCGAG CAGGCCGGCG GACCCTGCAC CTCATAGGCC 1200  
TCGCTGGCAT GGCGGGTTGT GCCATACTCA TGACCATCGC GCTAGCACTG CTGGAGCAGC 1260  
TACCTGGAT GTCTATCTG AGCATCGTGG CCATCTTTGG CTTTGTGGCC TTCTTTGAAG 1320  
TGGTCCCTGG CCCCATCCCA TGGTTCATCG TGGCTGAACT CTTCAGCCAG GGTCCACGTC 1380  
CAGCTGCCAT TGCCGTTGCA GGCTTCTCCA ACTGGACCTC AAATTTTCA TTTGGCATGT 1440  
40 GCTTCCAGTA TGTGGAGCAA CTGTGTGCTC CTAATGCTT CATCATCTT ACTGTGCTCC 1500  
TGGTTCGTT TTTCATCTTC ACCTACTTCA AAGTTCCTGA GACTAAAGGC CGGACCTTCG 1560  
ATGAGATCGC TTCCGGCTTC CGGCAGGGGG GAGCCAGCCA AAGTGATAAG ACACCCGAGG 1620  
AGCTGTTCCA TCCCCTGGGG GCTGATTCCC AAGTGTGAGT CGCCCCAGAT CACCAGCCCG 1680  
GCCTGCTCCC AGCAGCCCTA AGGATCTCTC AGGAGCACAG GCAGCTGGAT GAGACTTCCA 1740  
45 AACCTGACAG ATGTCAGCCG AGCCGGGCTC GGGGCTCCTT TCTCCAGCCA GCAATGATGT 1800  
CCAGAAGAAT ATTCAGGACT TAACGGCTCC AGGATTTTAA CAAAAGCAAG ACTGTTGCTC 1860  
AAATCTATT AGACAAGCAA CAGGTTTTAT AATTTTTTTA TTAATGATTT TGTATTTTTT 1920  
ATATCAGCCT GAGTCTCCTG TGCCACATC CCAGGCTTCA CCCTGAATGG TTCCATGCCT 1980  
GAGGGTGGAG ACTAAGCCCT GTCGAGACAC TTGCCTTCTT CACCCAGCTA ATCTGTAGGG 2040  
50 CTGGACCTAT GTCTAAGGA CACACTAATC GAACTATGAA CTACAAAGCT TCTATCCCAG 2100  
GAGGTGGCTA TGGCCACCCG TTCTGCTGGC CTGGATCTCC CCACTCTAGG GGTGAGGCTC 2160  
CATTAGGATT TGCCCTTCC CATCTCTTCC TACCCAAACA CTCAAATTAA TCTTTCTTTA 2220  
CCTGAGACCA GTTGGGAGCA CTGGAGTGGC GGGAGGAGAG GGGAGGGGCC AGTCTGGGCT 2280  
GCCGGGTTCT AGTCTCCTTT GCACTGAGGG CCACACTATT ACCATGAGAA GAGGGCCTGT 2340  
55 GGGAGCCTGC AAACCTACTG CTCAAGAAGA CATGGAGACT CCTGCCCTGT TGTGTATAGA 2400  
TGCAAGATAT TTATATATAT TTTTGGTTGT CAATATATAA TACAGACACT AAGTTATAGT 2460  
ATATCTGGAC AAGCCAACTT GTAAATACAC CACCTCACTC CTGTTACTTA CCTAAACAGA 2520  
TATAAATGGC TGGTTTTTAG AAACATGGTT TTGAAATGCT TGTGGATTGA GGGTAGGAGG 2580  
TTTGGATGGG AGTGAGACAG AAGTAAGTGG GGTGCAACC ACTGCAACGG CTTAGACTTC 2640  
60 GACTCAGGAT CCAGTCCCTT ACACGTACCT CTCATCAGTG TCCTCTTGCT CAAAAATCTG 2700  
TTTGATCCCT GTTACCCAGA GAATATATAC ATTCTTTATC TTGACATTCA AGGCATTTCT 2760  
ATCACATATT TGATAGTTGG TGTTCAAAAA AACACTAGTT TTGTGCCAGC CGTGATGCTC 2820  
AGGCTTGAAA TCGCATTATT TTGAATGTGA AGGGAA

Seq ID NO: 81 Protein sequence:  
Protein Accession #: NP\_006507.1

1 11 21 31 41 51  
| | | | | |  
70 MEPSSKKLTG RLMLAVGGAV LGSLLQFGYNT GVINAPQKVI EEFYNQTVWH RYGESILPTT 60  
LTTLSLSVA IFSVGMIGS FSVGLFVNRF GRRNSMLMMN LLAFVSAVLM GFSKLKKSFE 120  
MLILGRFIIG VYCGLTGTFV PMYVGEVSPT AFRGALGTLH QLGIVVGILI AQVFGLD SIM 180  
GNKDLWPLLL SIIFIPALLQ CIVLPFCPEP PRFLINRNE ENRAKSVLKK LRGTADVTHD 240  
75 LQEMKEESRQ MMREKKVTIL ELFRSPAYRQ PILIAVVLQL SQQLSGINAV FYYSTSIPEK 300  
AGVQPPVYAT IGSIVNTAF TVVSLFVVER AGRRTLHLIG LAGMAGCAIL MTIALALLEQ 360  
LPWMSYLSIV AIFGFVAFFE VPGPIPWFI VAEFLSQGPR PAIAVAGFS NWTSNFIVGM 420  
CFQYVEQLCG PYVFIIFTVL LVLFFIFTYF KVPETKGRTF DEIASGFRQG GASQSDKTPE 480  
ELFHPLGADS QV

Seq ID NO: 82 DNA sequence  
Nucleic Acid Accession #: BC001291  
Coding sequence: 44-541

1 11 21 31 41 51  
| | | | | |  
85 GGGGGCGCCG CGCGCTGACC CTCCCTGGGC ACCGCTGGGG ACGATGGCGC TGCTCGCCTT 60  
GCTGCTGGTC GTGGCCCTAC CGCGGGTGTG GACAGACGCC AACCTGACTG CGAGACAACG 120



	AGATCCAGAG	GACTCCCAGC	GAACGGACGA	GGGTGACAAT	AGAGTGTGGT	GTCATGTTTG	180
	TGAGAGAGAA	AACACTTTTCG	AGTGCCAGAA	CCCAAGGAGG	TGCAAATGGA	CAGAGCCATA	240
	CTGCGTTATA	GCGGCCGTGA	AAATATTTCC	ACGTTTTTTC	ATGGTTGCGA	AGCAGTGCTC	300
5	CGCTGGTTGT	GCAGCGATGG	AGAGACCCAA	GCCAGAGGAG	AAGCGGTTTC	TCCTGGAAGA	360
	GCCCATGCCC	TTCTTTTACC	TCAAAGTGTG	TAAAATTCGC	TACTGCAATT	TAGAGGGGCC	420
	ACCTATCAAC	TCATCAGTGT	TCAAAGAATA	TGCTGGGAGC	ATGGGTGAGA	GCTGTGGTGG	480
	GCTGTGGCTG	GCCATCCTCC	TGCTGCTGGC	CTCCATTGCA	GCCGGCCTCA	GCCTGTCTTG	540
	AGCCACGGGA	CTGCCACAGA	CTGAGCCTTC	CGGAGCATGG	ACTCGCTCCA	GACCGTTGTC	600
10	ACCTGTTGCA	TTAAACTTGT	TTTCTGTTGA	TTACCTCTTG	GTTTGACTTC	CCAGGGTCTT	660
	GGGATGGGAG	AGTGGGGATC	AGGTGCAGTT	GGCTCTTAAC	CCTCAAGGGT	TCTTTAACTC	720
	ACATTCAAG	GAAGTCCAGA	TCTCCTGAGT	AGTGATTTTG	GTGACAAGTT	TTTCTCTTTG	780
	AAATCAAACC	TTGTAACTCA	TTTATTGCTG	ATGGCCACTC	TTTTCCCTGA	CTCCCCTCTG	840
	CCTCTGAGGG	CTTCAGTATT	GATGGGGAGG	GAGGCCTAAG	TACCACTCAT	GGAGAGTATG	900
	TGCTGAGATG	CTTCCGACCT	TTCAGGTGAC	GCAGGAACAC	TGGGGGAGTC	TGAATGATTG	960
15	GGGTGAAGAC	ATCCCTGGAG	TGAAGGACTC	CTCAGCATGG	GGGGCAGTGG	GGCACACGTT	1020
	AGGGCTGCCC	CCATTCCAGT	GGTGGAGGCG	CTGTGGATGG	CTGCTTTTCC	TCAACCTTTC	1080
	CTACCAGATT	CCAGGAGGCA	GAAGATAACT	AATTGTGTTG	AAGAACTTA	GACTTCACCC	1140
	ACCAGCTGGC	ACAGGTGCAC	AGATTCTATA	ATTCCCACAC	GTGTGTGTTC	AACATCTGAA	1200
20	ACTTAGGCCA	AGTAGAGAGC	ATCAGGGTAA	ATGGCGTTCA	TTTCTCTGTT	AAGATGCAGC	1260
	CATCCATGGG	GAGCTGAGAA	ATCAGACTCA	AAGTTCACCC	AAAAACAAAT	ACAAGGGGAC	1320
	TTCAAAAGTT	CACGAAAAAA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA	AAA	

Seq ID NO: 83 Protein sequence:

Protein Accession #: AAH01291

25	1	11	21	31	41	51	
	MALLALLLVV	ALPRVWTDAN	LTARQRDPED	SQRTDEGDNR	VWCHVCEREN	TFECQNPERRC	60
30	KWTEPYCVIA	AVKIFPRFFM	VAKQCSAGCA	AMERPKPEEK	RFLLEEMPMPF	FYLKCKKIRY	120
	CNLEGPPIINS	SVFKEYAGSM	GESCGGLWLA	ILLLLASIAA	GLSLS		

Seq ID NO: 84 DNA sequence

Nucleic Acid Accession #: NM\_022893.1

Coding sequence: 229-2726

35	1	11	21	31	41	51	
	TTTTTTTTTT	TTTTTTGCTT	AAAAAAAAGC	CATGACGGCT	CTCCCACAAT	TCATCTTCCC	60
40	TGCGCCATCT	TTGTATTATT	TCTAATTAT	TTTGGATGTC	AAAAGGCACT	GATGAAGATA	120
	TTTTCTCTGG	AGTCTCCTTC	TTTCTAACCC	GGCTCTCCCG	ATGTGAACCG	AGCCGTCGTC	180
	CGCCCGCCGC	CGCCCGCCGC	GCCGCGCCGC	CCCGCCCCGC	AGCCCACCAT	GTCTCGCCGC	240
	AAGCAAGGCA	AACCCAGCA	CTTAAGCAAA	CGGGAATTCT	CGCCCGAGCC	TCTTGAAGCC	300
	ATTCTTACAG	ATGATGAACC	AGACCACGGC	CCGTGGGAG	CTCCAGAAGG	GGATCATGAC	360
45	CTCCTCACCT	GTGGGCAGTG	CCAGATGAAC	TTCCCATTTG	GGGACATTCT	TATTTTATC	420
	GAGCACAAAC	GGAAACAATG	CAATGGCAGC	CTCTGCTTAG	AAAAAGCTGT	GGATAAGCCA	480
	CCTTCCCTTT	CACCAATCGA	GATGAAAAAA	GCATCCAATC	CCGTGGAGGT	TGGCATCCAG	540
	GTACAGCCAG	AAGATGACGA	TGTGTTATCA	ACGTTCATCTA	GAAGAATTTG	CCCCAAACAG	600
	GAACACATAG	CAGATAAACT	TCTGCACCTG	AGGGGCCTCT	CCTCCCCTCG	TTCTGCACAT	660
50	GGAGCTCTAA	TCCCCACGCC	TGGGATGAGT	GCAGAATATG	CCCCGCAGGG	TATTTGTAAA	720
	GATGAGCCCA	GCAGCTACAC	ATGTACAACCT	TGCAAAACAGC	CATTACCCAG	TGCATGGTTT	780
	CTCTTGCAAC	ACGCACAGAA	CACTCATGGA	TTAAGAATCT	ACTTAGAAAG	CGAACACGGA	840
	AGTCCCCTGA	CCCCGCGGGT	TGGTATCCCT	TCAGGACTAG	GTGCAGAATG	TCCTTCCCAG	900
	CCACCTCTCC	ATGGGATTCA	TATTGCAGAT	AATAACCCCT	TTAACCTGCT	AAGAATACCA	960
55	GGATCAGTAT	CGAGAGAGGC	TTCGGGCTTG	GCAGAAGGGC	GCTTTCCACC	CACTCCCCCC	1020
	CTGTTTAGTC	CACCACCGAG	ACATCACTTG	GACCCCCACC	GCATAGAGCG	CCTGGGGGCG	1080
	GAAGAAATGG	CCCTGGCCAC	CCATCACCCG	AGTGCCCTTG	ACAGGGTGCT	GCGGTTGAAT	1140
	CCAATGGCTA	TGGAGCCTCC	CGCCATGGAT	TTCTCTAGGA	GACTTAGAGA	GCTGGCAGGG	1200
	AACACGTCTA	GCCCACCGCT	GTCCCCAGGC	CGGCCAGGCC	CTATGCAAAG	GTTACTGCAA	1260
60	CCATTCCAGC	CAGGTAGCAA	GCCGCCCTTC	CTGGCGAGCG	CCCCCTCCC	TCCTCTGCAA	1320
	TCCGCCCTC	CTCCCTCCCA	GCCCCCGGTC	AAGTCCAAGT	CATGCGAGTT	CTGCGGCAAG	1380
	ACGTTCAAAT	TTCAGAGCAA	CCTGGTGGTG	CACCGGCGCA	GCCACACGGG	CGAGAAGCCC	1440
	TACAAGTGCA	ACCTGTGCGA	CCACGCGTGC	ACCCAGGCCA	GCAAGCTGAA	GCGCCACATG	1500
	AAGACGCACA	TGCACAAATC	GTCCCCCATG	ACGGTCAAGT	CCGACGACGG	TCTCTCCACC	1560
65	GCCAGCTCCC	CGGAACCCGG	CACCAGCGAC	TTGGTGGGCA	GCGCCAGCAG	CGCGCTCAAG	1620
	TCCGTGGTGG	CCAAGTTCAA	GAGCGAGAAG	GACCCCAACC	TGATCCCGGA	GAACGGGGAC	1680
	GAGGAGGAAG	AGGAGGACGA	CGAGGAGAGG	GAAAGAGAGG	AGGAAGAGGA	GGAGGAGGAG	1740
	CTGACGGAGA	GCGAGAGGGT	GGACTACGGC	TTCGGGCTGA	GCCTGGAGGC	GGCGCGCCAC	1800
	CACGAGAACA	GCTCGCGGGG	CGCGGTCTGT	GGCGTGGGCG	ACGAGAGCCG	CGCCCTGCCC	1860
70	GACGTCATGC	AGGGCATGGT	GCTCAGCTCC	ATGCAGCACT	TCAGCGAGGC	CTTCCACCAG	1920
	GTCTTGGGCG	AGAAGCATAA	GCGCGGCCAC	CTGGCCGAGG	CCGAGGGCCA	CAGGGACACT	1980
	TGCGACGAAG	ACTCGGTGGC	CGGCGAGTCC	GACCGCATAG	ACGATGGCAC	TGTTAATGGC	2040
	CGCGGTGCT	CTCCGGGCGA	GTCGGCCTCG	CGGGCCCTGT	CCAAAAGCT	GCTGCTGGGC	2100
	AGCCCCAGCT	CGCTGAGCCC	CTTCTCTAAG	CGCATCAAGC	TCGAGAAGGA	GTTCTGACCTG	2160
75	CCCCCGGCCA	CGATGCCCAA	CACGGAGAAC	GTGTACTCGC	AGTGGCTCGC	CGGCTACGCG	2220
	GCCTCCAGGC	AGCTCAAAGA	TCCCTTCTCT	AGCTTCGGAG	ACTCCAGACA	ATCGCCTTTT	2280
	GCCTCCTCGT	CGGAGCACTC	CTCGGAGAAC	GGGAGCTTGC	GCTTCTCCAC	ACCGCCCGGG	2340
	GAGCTGGACG	GAGGGATCTC	GGGGCGCAGC	GGCACGGGAA	GTGGAGGGAG	CACGCCCAT	2400
	ATTAGTGGTC	CGGGCACGGG	CAGGCGCAGC	TCAAAAGAGG	GCAGACGCAG	CGACACTTGT	2460
80	GAGTACTGTG	GGAAAGTCTT	CAAGAAGTGT	AGCAATCTCA	CTGTCCACAG	GAGAAGCCAC	2520
	ACGGGCGAAA	GGCCTTATAA	ATGCGAGCTG	TGCAACTATG	CCTGTGCCCA	GAGTAGCAAG	2580
	CTCACCAGGC	ACATGAAAAC	GCATGGCCAG	GTGGGGAAGG	ACGTTTACAA	ATGTGAAATT	2640
	TGTAAGATGC	CTTTTAGCGT	GTACAGTACC	CTGGAGAAAC	ACATGAAAAA	ATGGCACAGT	2700
	GATCGAGTGT	TGAATAATGA	TATAAAAACT	GAATAGAGGT	ATATTAATAC	CCCTCCCTCA	2760
85	CTCCACCTG	ACACCCCTTT	TTCCACCTTT	CCCTTCCCC	ATCGCCCTCC	AGCCCACTC	2820
	CCTGTAGGAT	TTTTTTCTAG	TCCCATGTGA	TTTAAACAAA	CAAACAAACA	AACAGAAGTA	2880
	ACGAAGCTAA	GAATATGAGA	GTGCTTGTCA	CCAGCACACC	TGTTTTTTTT	CTTTTCTTTT	2940
	TTCTTTTTTC	TTTTTCCTTT	TTTTTTTTTT	TCCTTTATGT	TCTCACCGTT	TGAATGCATG	3000



	ATCTGTATGG	GGCAATACTA	TTGCATTTTA	CGCAAACCTT	GAGCCTTTCT	CTTGTGCAAT	3060
	AATTTACATG	TTGTGTATGT	TTTTTTTAA	ACTTAGACAG	CATGTATGGT	ATGTTATGGC	3120
	TATTTTAAAT	TGTCCCTAAT	TCGTTGCTGA	GCAAAACATG	TGCTGTTTCC	AGTTCCGTTT	3180
5	TGAGAGAAAA	AGAGAGAGAG	AGAGAAAAAG	ACCATGCTGC	ATACATTCTG	TAATACATAT	3240
	CATGTACAGT	TTTATTTTAT	AACGTGAGGA	GGAAAAACAG	TCTTTGGATT	AACCCTCTAT	3300
	AGACAGAATA	GATAGCACTG	AAAAAAAATC	TCTATGAGCT	AAATGTCTGT	CTCTAAAGGG	3360
	TTAAATGTAT	CAATTGGAAA	GGAAGAAAAA	AGGCCTTGAA	TTGACAAAT	AACAGAAAAA	3420
	CAGAACAAGT	TTATTCTATC	ATTTGGTTTT	AAAATATGAG	TGCCTTGGAT	CTATTTAAAC	3480
10	CACATCGATG	GTTCTTTCTA	CTTGTTATAA	ACTTGTAGCT	TAATTCAGCA	TTGGGTGAGG	3540
	TAATAAACCT	TAGGAACCTAG	CATATAATTC	TATATTGTAT	TTCTCACAAC	AATGGCTACC	3600
	TAAAAAGATG	ACCCATTATG	TCCTAGTTAA	TCATCATTTT	TCCTTTAGTT	TAATTTTATA	3660
	AACAAAACCTG	ATTATACCAG	TATAAAAGCT	ACTTTGCTCC	TGGTGAGAGC	TTAAAAGAAA	3720
	TGGGCTGTTT	TGCCCAAAGT	TTTATTTTTT	TTAAACAATG	ATTAAATTGA	ATGTGTAATG	3780
15	TGCAAAAGCC	CTGGAACGCA	ATTAAATACA	CTAGTAAGGA	GTTTCATTTA	TGAAGATATT	3840
	TGCTTTAATA	ATGTCTTTTT	AAAAATACATG	GCACCAAAAG	AAATAGATCC	AGATCTACTT	3900
	GGTTGTCAAG	TGGACAATCA	AATGATAAAC	TTAAGACCT	TGTATACCAT	ATTGAAAGGA	3960
	AGAGGCTGAC	AATAAGGTTT	GACAGAGGGG	AACAGAAGAA	AATAATATGA	TTTATTAGCA	4020
	CAACGTGGTA	CTATTTGCCA	TTTAAACTA	GAACAGGTAT	ATAAGCTAAT	ATTGATACAA	4080
20	TGATGATTAA	CTATGAATTC	TTAAGACTTG	CATTTAAATG	TGACATTCTT	AAAAAAGAA	4140
	GAGAAAGAAT	TTTAAGAGTA	GCAGTATATA	TGCTGTGCT	CCCTAAAAGT	TGTACTTCAT	4200
	TTCTTTTCCA	TACACTGTGT	GCTATTTGTG	TTAACATGGA	AGAGGATTCA	TTGTTTAT	4260
	TTTTATTTTT	TAAATTTTTT	CTTTTTTATT	AAGCTAGCAT	CTGCCCCAGT	TGGTGTTCAT	4320
	ATAGCACTTG	ACTCTGCCTG	TGATATCTGT	ATCTTTTCTC	TAATCAGAGA	TACAGAGGTT	4380
25	GAGTATAAAA	TAAACCTGCT	CAGATAGGAC	AATTAAGTGC	ACTGTACAAT	TTTCCAGTT	4440
	TACAGGTCTA	TACTTAAGGG	AAAAGTTGCA	AGAATGCTGA	AAAAAATTG	AACACAATCT	4500
	CATTGAGGAG	CATTTTTTAA	AAACTAAAAA	AAAAAAACT	TTGCCAGCCA	TTTACTTGAC	4560
	TATTGAGCTT	ACTTACTTGG	ACGCAACATT	GCAAGCGCTG	TGAATGGAAA	CAGAATACAC	4620
	TTAACATAGA	AATGAATGAT	TGCTTTCGCT	TCTACAGTGC	AAGGATTTTT	TTGTACAAAA	4680
30	CTTTTTTAAA	TATAAATGTT	AAGAAAAATT	TTTTTTTAAA	AACACTTCAT	TATGTTTAGG	4740
	GGGGAACCTG	ATTTTAGGGT	TCCATTGTCT	TGGTGGTGTT	ACAAGACTTG	TTATCCATTT	4800
	AAAAATGGTA	GTGGAAATTC	TATGCCTTGG	ATACACACCG	CTCTTCAGGT	TGTAAAAAAA	4860
	AAAAACATAC	ATTGGGGAAA	GGTTTAAGAT	TATATAGTAC	TTAAATATAG	GAAATGCAC	4920
	ACTCATGTTG	ATTCTATGCT	TAAATACAT	TTATGGTCTT	TTTTCTGTAT	TTCTAGAATG	4980
35	GTATTTGAAT	TAAATGTTCA	TCTAGTGTTA	GCACTATAG	TATTTATATT	GAAGCTTGTA	5040
	TTTTTAACTG	TTGCTTGTTT	TCTTAAAGG	TATCAATGTA	CCTTTTTTGG	TAGTGAAAAA	5100
	AAAAAAGACA	GGCTGCCACA	GTATATTTTT	TTAATTTGGC	AGGATAATAT	AGTGCAAATT	5160
	ATTTGTATGC	TTCAAAAAAA	AAAAAAGAG	AGAAACAAAA	AAGTGTGACA	TTACAGATGA	5220
	GAAAGCCATAT	AATGGCGGTT	TGGGGGAGCC	TGCTAGAATG	TCACATGGAT	GGCTGTCATA	5280
40	GGGGTTGTAC	ATATCCTTTT	TTGTTCCCTT	TTCTGCTGCT	CATACTGTAT	GCAGTACTGC	5340
	AAGCTAATAA	CGTTGGTTTG	TTATGTAGTT	TGCTTTTGTG	CCCTTTCCTT	CTATCACCTT	5400
	ACATTCAGAG	ATCTTACCTT	CATATGCAGT	AAAAAGAAAG	AAGAAAAAAA	AAGGAAAAAA	5460
	AAAAAAAAC	CAATGTTTTG	CAGTTTTTTT	CATTGCCAAA	AACTAAATGG	TGCTTTATAT	5520
	TTAGATTGGA	AAGAATTTCA	TATGCAAAGC	ATATTAAAGA	GAAAGCCCGC	TTTAGTCAAT	5580
45	ACTTTTTTGT	AAATGGCAAT	GCAGAATATT	TTGTTATTGG	CCTTTTCTAT	TCCTGTAATG	5640
	AAAGCTGTTT	GTCGTAACCT	GAAATTTTAT	CTTTTACTAT	GGGAGTCACT	ATTTATTATT	5700
	GCTTATGTGC	CCTGTTCAAA	ACAGAGGCAT	TTAATTTGAT	CTTTTATTTT	TCTTTGTTTT	5760
	TATTTTTTTT	TTTATTTAGA	TGACCAAGG	TCAATACAAC	CTGGCTTTTT	ATTGTATTTG	5820
	TTTCTGGTCT	TTGTTAAGTT	CTATTGGAAA	AACCACTGTC	TGTGTTTTTT	TGGCAGTTGT	5880
50	CTGCATTAAC	CTGTTTCATAC	ACCCATTTTG	TCCCTTTATT	GAAAAAATAA	AAAAAATTAA	5940
	A						

Seq ID NO: 85 Protein sequence:  
Protein Accession #: NP\_075044.1

55	1	11	21	31	41	51	
	MSRRKQGKPKQ	HLSKREFSPE	PLEAILTDDE	PDHGPLGAPE	GDHDLTTCGQ	CQMNFPPLGDI	60
	LIFIEHKRKQ	CNGSLCLEKA	VDKPPSPSPI	EMKKASNPVE	VGIQVTPEDD	DCLSTSSRRI	120
60	CPKQEHADK	LLHWRGLSSP	RSAGALIPQ	PGMSAEYAPQ	GICKDEPSSY	TCTTCKQPF	180
	SAWFLQLQHAQ	NTHGLRIYLE	SEHGSPPLTPR	VGIPSGLGAE	CPSQPPLHGI	HIADNPFNL	240
	LRIPGSVSRE	ASGLAEGFRP	PTPPLFSPPP	RHHLDPHRIE	RLGAEEMALA	THHPSAFDRV	300
	LRLNPMAMEP	PAMDFSRRRLR	ELAGNTSSPP	LSPGRPSPMQ	RLLQPFQPGS	KPPFLATPPL	360
	PPLQSAAPPSP	QPPVKSKSCE	FCGKTFKFQS	NLVVHRRSHT	GEKPYKCNLC	DHACTQASKL	420
65	KRHKMTHMHK	SSPMTVKSD	GLSTASSPEP	GTSDLVGSAS	SALKSVVAKF	KSENDPNLIP	480
	ENGDEEEED	DEEEEEEED	EEELTESER	VDYGFGLSLE	AARHHENSSR	GAVVGVGDES	540
	RALPDVMOQM	VLSSMQHFSE	AFHQVLGEKH	KRGHLAEAE	HRDTCDEDSV	AGESDRIDDG	600
	TVNGRGCSFG	ESASGGLSKK	LLLGSPSSLS	PFSKRIKLEK	EFDLPPATMP	NTEVYSQWL	660
	AGYAASRQLK	DPFLSFGDSR	QSPFASSSEH	SENGSLRFS	TPPGELDGGI	SGRSGTSGSG	720
70	STPHISGPGT	GRPSSKEGRR	SDTCEYCGKV	FKNCSNLTVH	RRSHTGERPY	KCELCNYACA	780
	QSSKLTRHMK	THGQVGKDVY	KCEICKMPFS	VYSTLEKHMK	KWHSRDLN	DIKTE	

Seq ID NO: 86 DNA sequence  
Nucleic Acid Accession #: XM\_035292.2  
Coding sequence: 53-1576

75	1	11	21	31	41	51	
80	GCTCGCTGGG	CCGCGGCTCC	CGGGTGTCCC	AGGCCCGGCC	GGTGCAGAGA	GCATGGCGGG	60
	TGCGGGCCCG	AAGCGGCGCG	CGCTAGCGGC	GCCGGCGGCC	GAGGAGAAGG	AAGAGCGCG	120
	GGAGAAGATG	CTGGCCGCCA	AGAGCGCGGA	CGGCTCGGCG	CCGGCAGGCG	AGGGCGAGGG	180
	CGTGACCTTG	CAGCGGAACA	TCACGCTGCT	CAACGGCGTG	GCCATCATCG	TGGGGACCAT	240
	TATCGGCTCG	GGCATCTTCG	TGACGCCAC	GGGCGTGCTC	AAGGAGGCAG	GCTCGCCGGG	300
	GCTGGCGCTG	GTGGTGTGGG	CCGCTGTGGG	CGTCTTCTCC	ATCGTGGGCG	CGCTCTGCTA	360
85	CGCGGAGCTC	GGCACCACCA	TCTCCAATC	GGCGGCGGAC	TACGCCTACA	TGCTGGAGGT	420
	CTACGGCTCG	CTGCCCGCCT	TCCTCAAGCT	CTGGATCGAG	CTGCTCATCA	TCCGGCCTTC	480
	ATCGCAGTAC	ATCGTGGCCC	TGGTCTTCGC	CACCTACCTG	CTCAAGCCGC	TCTTCCCCAC	540



5 CTGCCCCGTG CCCGAGGAGG CAGCCAAGCT CGTGGCCTGC CTCTGCGTGC TGCTGCTCAC 600  
 GGCCGTGAAC TGCTACAGCG TGAAGGCCCG CACCCGGGTC CAGGATGCCT TTGCCGCCGC 660  
 CAAGCTCCTG GCCCTGGCCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GGAAGGGTGA 720  
 TGTGTCCAAT CTAGATCCCA ACTTCTCATT TGAAGGCACC AAAGTGGATG TGGGGAACAT 780  
 TGTGCTGGCA TTATACAGCG GCCTCTTTGC CTATGGAGGA TGGAATTACT TGAATTTCTG 840  
 CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCCCTG GCCATCATCA TCTCCCTGCC 900  
 CATCGTGACG CTGGTGTACG TGCTGACCAA CCTGGCCTAC TTCACCACCC TGTCCACCGA 960  
 GCAGATGCTG TCGTCCGAGG CCGTGGCCGT GGAATTCGGG AACTATCACC TGGGCGTCAT 1020  
 10 GTCTTGGATC ATCCCGCTCT TCGTGGGCTT GTCCTGCTTC GGCTCCGTCA ATGGGTCCCT 1080  
 GTTACATCC TCCAGGCTCT TCTTCGTGGG GTCCCGGGA GGCCACCTGC CCTCCATCCT 1140  
 CTCCATGATC CACCCACAGC TCCTCACCCC CGTGCCGTCC CTCGTGTTCA CGTGTGTGAT 1200  
 GACGCTGCTC TACGCCTTCT CCAAGGACAT CTTCTCCGTC ATCAACTTCT TCAGCTTCTT 1260  
 CAACTGGCTC TCGTGGGCC TGGCCATCAT CCGCATGATC TGGCTGCGCC ACAGAAAGCC 1320  
 TGAGCTTGAG CGGCCATCA AGGTGAACCT GGCCCTGCCT GTGTCTTCA TCCTGGCCTG 1380  
 15 CCTCTTCTCT CTTCTCGTCT CTTCTCGTCT GACACCCGTG GAGTGTGGCA TCGGCTTCAC 1440  
 CATCATCCTC AGCGGGCTGC CCGTCTACTT CTTGCGGGTC TGGTGGAAAA ACAAGCCCAA 1500  
 GTGGCTCCTC CAGGGCATCT TCTCCACGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGGT 1560  
 CCCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

20 Seq ID NO: 87 Protein sequence:  
 Protein Accession #: XP\_035292.2

25 1 11 21 31 41 51  
 | | | | |  
 MAGAGPKRRA LAAPAAEKE EAREKMLAAK SADGSAPAGE GEGVTLQRNI TLLNGVAIIV 60  
 GTIIGSGIFV TPTGVLKEAG SPGLALVWA ACGVFSIVGA LCYaelGTTI SKSGGDYAYM 120  
 LEVYGSIPAF LKLWIELLII RPSSQYIVAL VFATYLLKPL FPTCPVPEEA AKLVACLCLV 180  
 LLTAVNCYSV KAATRVQDAF AAALLLALAL IILLGFVQIG KGDVSNLDPN FSEGTCLDV 240  
 30 GNIVLALYSG LFAYGGWNYL NFVTEEMINP YRNLPLAIII SLPIVTLVYV LTNLAYFTTL 300  
 STEQMLSEA VAVDFGNYHL GVMSWIIPVF VGLSCFGSVN GSLFTSSRLF FVGSREGHLP 360  
 SILSMIHPQL LTPVPSLVFT CVMTLLYAFS KDIFSVINFF SFFNWLCVAL AIIGMIWLRH 420  
 RKPELERPIK VNLALPVFFI LACFLIAVS FWKTPVECGI GFTIILSGLP VYFFGVWVKN 480  
 KPKWLLQGIF STTVLCQKLM QVVPQET

35 Seq ID NO: 88 DNA sequence  
 Nucleic Acid Accession #: NM\_005268.1  
 Coding sequence: 168-989

40 1 11 21 31 41 51  
 | | | | |  
 TAAAAAGCAA AAGAATTCGC GGCCGCGTCG ACACGGGCTT CCCCAGAAAC CTTCCCCGCT 60  
 TCTGGATATG AAATTCAAGC TGCTTGCTGA GTCCTATTGC CGGCTGCTGG GAGCCAGGAG 120  
 AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACGCGTGG GTCCACCATG AACTGGAGTA 180  
 45 TCTTTGAGGG ACTCCTGAGT GGGGTCAACA AGTACTCCAC AGCCTTTGGG CGCATCTGGC 240  
 TGTCTCTGGT CTTTCATCTT CCGTGCTGG TGTACCTGGT GACGGCCGAG CGTGTGTGGA 300  
 GTGATGACCA CAAGGACTTC GACTGCAATA CTCGCCAGCC CGGCTGCTCC AACGTCTGCT 360  
 TTGATGAGTT CTTCCCTGTG TCCCATGTGC GCCTCTGGGC CCTGCAGCTT ATCCTGGTGA 420  
 CATGCCCTC ACTGCTCGTG GTCATGCACG TGGCCTACCG GGAGGTTGAG GAGAAGAGGC 480  
 50 ACCGAGAAGC CCATGGGGAG AACAGTGGGC GCCTCTACCT GAACCCCGGC AAGAAGCGGG 540  
 GTGGGCTCTG GTGGACATAT GTCTGCAGCC TAGTGTTCAG GGCGAGCGTG GACATCGCCT 600  
 TTCTCTATGT GTTCCACTCA TTCTACCCCA AATATATCCT CCCTCCTGTG GTCAAGTGCC 660  
 ACGCAGATCC ATGTCCCAAT ATAGTGGACT GCTTCATCTC CAAGCCCTCA GAGAAGAACA 720  
 TTTTCACCCCT CTTTCATGGT GCCACAGCTG CCATCTGCAT CCTGCTCAAC CTCGTGGAGC 780  
 55 TCATCTACCT GGTGAGCAAG AGATGCCACG AGTGCCCTGG AGCAAGGAAA GCTCAAGCCA 840  
 TGTGCACAGG TCATCACCCC CACGGTACCA CCTCTTCTG CAAACAAGAC GACCTCCTTT 900  
 CGGGTGACCT CATCTTTCTG GGCTCAGACA GTCATCCTCC TCTCTTACCA GACCGCCCCC 960  
 GAGACCATGT GAAGAAAACC ATCTTGTGAG GGGCTGCCTG GACTGGTCTG GCAGGTTGGG 1020  
 CCTGGATGGG GAGGCTCTAG CATCTCTCAT AGTGTCAACC TGAGAGTGGG GGAGCTAAGC 1080  
 60 CATGAGGTAG GGGCAGGCAA GATGAGGAT TCAGACGCTC TGGGAGCCAG TTCCTAGTCC 1140  
 TCAACTCCAG CCACCTGCCC CAGCTCGACG GCACTGGGCC AGTTCCCCCT CTGCTCTGCA 1200  
 GCTCGGTTTC CTTTTCTAGA ATGGAAATAG TGAGGGCCAA TGC

65 Seq ID NO: 89 Protein sequence:  
 Protein Accession #: NP\_005259.1

70 1 11 21 31 41 51  
 | | | | |  
 MNWSIFEGLL SGVNKYSTAF GRIWLSLVFI FRVLVYLVT A ERVWSDDHKD FDCNTRQPGC 60  
 SNVCFDEFFP VSHVRLWALQ LILVTCPSLL VVMHVAAYREV QEKHRHREAHG ENSGRLYLNP 120  
 GKRGGLWWT YVCSLVFKAS VDIALLYVFH SFYPKYILPP VVKCHADPCP NIVDCFISKP 180  
 SEKNIFTLEF VATAAICILL NLVELIYLV KRCHECLAAR KAQAMCTGHH PHGTTSSCKQ 240  
 DDLLSGDLIF LGSDSHPPLL PDRPRDHVKK TIL

75 Seq ID NO: 90 DNA sequence  
 Nucleic Acid Accession #: NM\_002391.1  
 Coding sequence: 26-457

80 1 11 21 31 41 51  
 | | | | |  
 CGGGCGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCCTCCTCC TCACCCTCCT 60  
 CGCCCTGCTG GCGCTCACCT CCGCGGTCGC CAAAAAGAAA GATAAGGTGA AGAAGGGCGG 120  
 CCCGGGAGC GAGTGCCTG AGTGGGCTGC GGGGCCCTGC ACCCCAGCA GCAAGGATTG 180  
 85 CGGCGTGGGT TTCCGCGAGG GCACCTGCGG GGCCAGACC CAGCGCATCC GGTGCAGGGT 240  
 GCCCTGCAAC TGGAAGAAGG AGTTTGGAGC GCACTGCAAG TACAAGTTTG AGAACTGGGG 300  
 TGCGTGTGAT GGGGACACAG GCACCAAAGT CCGCCAAGGC ACCCTGAAGA AGGCGCGCTA 360



CAATGCTCAG TGCCAGGAGA CCATCCGCGT CACCAAGCCC TGCACCCCCA AGACCAAAGC 420  
 AAAGGCCAAA GCCAAGAAAG GGAAGGGAAA GGACTAGACG CCAAGCCTGG ATGCCAAGGA 480  
 GCCCCTGGTG TCACATGGGG CCTGGCCACG CCCTCCTCT CCCAGGCCCG AGATGTGACC 540  
 CACCAAGTGC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCCTGCCT TGTCCCTCTC 600  
 ACTCCCCAGC CCCACCCCTA AGTGCCCAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT 660  
 TGAGCCTCCC CCAAAGCAAT GTGAGTCCCA GAGCCCGCTT TTGTTCTTCC CCACAATTCC 720  
 ATTACTAAGA AACACATCAA ATAAACTGAC TTTTCCCCC CAATAAAAGC TCTTCTTTT 780  
 TAATAT

Seq ID NO: 91 Protein sequence:  
 Protein Accession #: NP\_002382.1

1 11 21 31 41 51  
 | | | | |  
 15 MQHRGFLLLT LLALLALTSA VAKKKDKVKK GPGSECAEW AWGPCTPSSK DCGVGFREGT 60  
 CGAQTQRIRC RVPCNWKKEF GADCKYKFEN WGACDGGTGT KVRQGTLLKA RYNAQCQETI 120  
 RVTKPCTPKT KAKAKAKKGK GKD

Seq ID NO: 92 DNA sequence  
 Nucleic Acid Accession #: NM\_005130.1  
 Coding sequence: 98-802

1 11 21 31 41 51  
 | | | | |  
 25 CTCTACCTGA CACAGCTGCA GCCTGCAATT CACTCCCACT GCCTGGGATT GCACTGGATC 60  
 CGTGTGCTCA GAACAAGGTG AACGCCACG TGCAGCCATG AAGATCTGTA GCCTCACCTT 120  
 GCTCTCCTTC CTCCTACTGG CTGCTCAGGT GCTCCTGGTG GAGGGGAAAA AAAAAGTGAA 180  
 30 GAATGGACTT CACAGCAAAG TGGTCTCAGA ACAAAGGAC ACTCTGGGCA ACACCCAGAT 240  
 TAAGCAGAAA AGCAGGCCCG GGAACAAAG CAAGTTTGTG ACCAAAGACC AAGCCAACTG 300  
 CAGATGGGCT GCTACTGAGC AGGAGGAGGG CATCTCTCTC AAGGTTGAGT GCACTCAATT 360  
 GGACCATGAA TTTTCTGTG TCTTTGCTGG CAATCCAACC TCATGCCTAA AGCTCAAGGA 420  
 TGAGAGAGTC TATTGGAAAC AAGTTGCCCG GAATCTGCGC TCACAGAAAG ACATCTGTAG 480  
 35 ATATTCCAAG ACAGCTGTGA AAACCAGAGT GTGCAGAAAG GATTTTCCAG AATCCAGTCT 540  
 TAAGCTAGTC AGCTCCACTC TATTTGGGAA CACAAAGCCC AGGAAGGAGA AAACAGAGAT 600  
 GTCCCCCAGG GAGCACATCA AGGGCAAAGA GACCACCCCT TCTAGCCTAG CAGTGACCCA 660  
 GACCATGGCC ACCAAAGCTC CCGAGTGTGT GGAGGACCCA GATATGGCAA ACCAGAGGAA 720  
 GACTGCCCTG GAGTTCTGTG GAGAGACTTG GAGCTCTCTC TGCACATTCT TCCTCAGCAT 780  
 40 AGTGCAGGAC ACGTCATGCT AATGAGGTCA AAAGAGAACG GGTTCTTTTA AGAGATGTCA 840  
 TGTCTTAAGT CCCTCTGTAT ACTTTAAAGC TCTCTACAGT CCCCCAAAA TATGAACTTT 900  
 TGTGCTTAGT GAGTGCAACG AAATATTTAA ACAAGTTTTC TATTTTTTGC TTTTGTGTTT 960  
 TGGAATTTGC CTTATTTTTC TTGGATGCGA TGTTGAGAGG CTGTTTCTCT CAGCATGTAT 1020  
 TTCCATGGCC CACACAGCTA TGTGTTTGTG CAGCGAAGAG TCTTTGAGCT GAATGAGCCA 1080  
 45 GAGTGATAAT TTCAGTGCAA CGAAGTTTCT GCTGAATTAA TGGTAATAAA ACTCTGGGTG 1140  
 TTTTTCAAAA AAAAAAAAAA AAA

Seq ID NO: 93 Protein sequence:  
 Protein Accession #: NP\_005121.1

1 11 21 31 41 51  
 | | | | |  
 50 MKICSLTLLS FLILLAAQVLL VEGKKKVKNG LHSKVSEQK DTLGNTQIKQ KSRPGNKGKF 60  
 VTKDQANCRW AATEQEEGIS LKVECTQLDH EFSCVFAGNP TSCLKLKDER VYWKQVARNL 120  
 55 RSQKIDICRY KTAVKTRVCR KDFPESSLKL VSSTLFGNTK PRKEKTEMSP REHIKGETT 180  
 PSSLAVTQTM ATKAPECVED PDMANQRKTA LEFCGETWSS LCTFFLSIVQ DTSC

Seq ID NO: 94 DNA sequence  
 Nucleic Acid Accession #: NM\_012101  
 Coding sequence: 125-1891

1 11 21 31 41 51  
 | | | | |  
 60 CTCCTCACAG GTGTGTCTCT AGTCCTCGTG GTTGCCTGCC CCACTCCCTG CCGAGACGCC 60  
 TGCCAGAAAG GTCACCTATC CTGAACCCCA GCAAGCCTGA AACAGCTCAG CCAAGCACCC 120  
 65 TCGATGGAA GCTGCAGATG CCTCCAGAG CAACGGGTCTG AGCCAGAAAG CCAGGGATGC 180  
 CCGGAGCCCG TCGGGCCCCA GTGGCAGCCT GGAGAATGGC ACCAAGGCTG ACGGCAAGGA 240  
 TGCCAGAGAC ACCAACGGGC ACGGCGGGGA GGCAGCTGAG GGCAAGAGCC TGGGCAGCGC 300  
 CCTGAAGCCA GGGGAAGGTA GGAGCGCCCT GTTCGCGGGC AATGAGTGGC GGCAGCCCAT 360  
 70 CATCCAGTTT GTCGAGTCCG GGGACGACAA GAACTCCAAC TACTTCAGCA TGGACTCTAT 420  
 GGAAGGCAAG AGGTGCGCGT ACGCAGGGCT CCAGCTGGGG GCTGCCAAGA AGCCACCCGT 480  
 TACCTTTGCC GAAAAGGGCG ACGTGCAGCA GTCCATTTTC TCGGAGTCCC GGAAGCCAC 540  
 GGTGTCCATC ATGGAGCCCG GGGAGACCCG GCGGAACAGC TACCCCGGG CCGACACGGG 600  
 CCTTTTTC CAAGTCAAGT CCGGCTCCGA GGAGGTGCTG TGCGACTCCT GCATCGGCAA 660  
 75 CAAGCAGAAG GCGGTCAAGT CCTGCCTGGT GTGCCAGGCC TCCTTCTGCG AGCTGCATCT 720  
 CAAGCCCCAC CTGGAGGGCG CCGCCTTCCG AGACCACAG CTGCTCGAGC CCATCCGGGA 780  
 CTTTGAGGCC CGCAAGTGTC CCGTGCATGG CAAGACGATG GAGCTCTTCT GCCAGACCGA 840  
 CCAGACCTGC ATCTGCTACC TTTGCATGTT CCAGGAGCAC AAGAATCATA GCACCGTGAC 900  
 AGTGGAGGAG GCCAAGGCCG AGAAGGAGT GGAGCTGTCA CTGCAAAAGG AGCAGCTGCA 960  
 80 GCTCAAGATC ATTGAGATTG AGGATGAAGC TGAGAAGTGG CAGAAGGAGA AGGACCGCAT 1020  
 CAAGAGCTTC ACCACCAATG AGAAGGCCAT CCTGGAGCAG AACTTCCGGG ACCTGGTGCG 1080  
 GGACCTGGAG AAGCAAAAGG AGGAAGTGAG GGCTGCGCTG GAGCAGCGGG AGCAGGATGC 1140  
 TGTGGACCAA GTGAAGGTGA TCATGGATGC TCTGGATGAG AGAGCCAAGG TGCTGCATGA 1200  
 GGACAAGCAG ACCCGGGAGC AGCTGCTATG CATCAGCGAC TCTGTGTTGT TTCTGCAGGA 1260  
 85 ATTTGGTGCA TTGATGAGCA ATTACTCTCT CCCCCACCT CTGCCCACCT ATCATGTCCT 1320  
 GCTGGAGGGG GAGGGCCTGG CACAGTCACT AGGCAACTTC AAGGACGACC TGCTCAATGT 1380  
 ATGCATGCGC CACGTTGAGA AGATGTGCAA GGCGGACCTG AGCCGTAAC TATTGAGAG 1440  
 GAACCACATG GAGAACGGTG GTGACCATCG CTATGTGAAC AACTACACGA ACAGCTTCGG 1500



5 GGGTGAGTGG AGTGCAACGG ACACCATGAA GAGATACTCC ATGTACCTGA CACCCAAAGG 1560  
 TGGGGTCCGG ACATCATACC AGCCCTCGTC TCCTGGCCGC TTCACCAAGG AGACCACCCA 1620  
 GAAGAATTTT AACAATCTCT ATGGCACCAA AGGTAACCTA ACCTCCCGGG TCTGGGAGTA 1680  
 CTCCTCCAGC ATTCAGAACT CTGACAATGA CCTGCCCGTC GTCCAAGGCA GCTCCTCCTT 1740  
 CTCCCTGAAA GGCTATCCCT CCCTCATGCG GAGCCAAAGC CCAAGGCCG AGCCCCAGAC 1800  
 TTGGAAATCT GGCAAGCAGA CTATGCTGTC TCACTACCGG CCATTCTACG TCAACAAAGG 1860  
 CAACGGGATT GGGTCCAACG AAGCCCCATG AGCTCCTGGC GGAAGGAACG AGGCGCCACA 1920  
 CCCCTGCTCT TCCTCCTGAC CCTGCTGCTC TTGCCCTCTA AGCTACTGTG CTTGTCTGGG 1980  
 10 TGGGAGGGAG CCTGGTCCCTG CACCTGCCCT CTGCAGCCCT CTGCCAGCCT CTTGGGGGCA 2040  
 GTTCCGGCCT CTCCGACTTC CCCACTGGCC AACTCCATT CAGACTCCTT TCCTGCCTTG 2100  
 TGACCTCAGA TGGTCACCAT CATTCTGTG CTCAGAGGCC AACCCTACAG AGGGGTGAGA 2160  
 TAGGTTGGGG CCTGCCCTAA CCCGCCAGCC TCCTCCTCTC GGGCTGGATC TGGGGGCTAG 2220  
 CAGTGAGTAC CCGCATGGTA TCAGCCTGCC TCTCCCGCCC ACGCCCTGCT GTCTCCAGGC 2280  
 CTATAGACGT TTCTCTCCAA GGCCCTATCC CCCAATGTTG TCAGCAGATG CCTGGACAGC 2340  
 15 ACAGCCACCC ATCTCCCATT CACATGGCCC ACCTCCTGCT TCCAGAGGA CTGGCCCTAC 2400  
 GTGCTCTCTC TCGTCTTACC TATCAATGCC CAGCATGGCA GAACCTGCAG TGGCCAAGGG 2460  
 CTGCAGATGG AAACCTCTCA GTGTCTTGAC ATCACCCTAC CCAGGCGGTG GGTCTCCACC 2520  
 ACAGCCACTT TGAGTCTGTG GTCCCTGGAG GGTGGCTTCT CTGACTGGC AGGATGACCT 2580  
 TAGCCAAGAT ATTCTCTGT TCCCTCTGCT GAGATAAAGA ATTCCCTTAA CATGATATAA 2640  
 20 TCCACCCATG CAAATAGCTA CTGGCCAGC TACCATTAC CATTGCTTA CAGAATTTCA 2700  
 TTCAGTCTAC ACTTTGGCAT TCTCTCTGGC GATGGAGTGT GGCTGGGCTG ACCGCAAAAG 2760  
 GTGCCCTTACA CACTGCCCCC ACCCTCAGCC GTTGCCCCAT CAGAGGCTGC CTCTCTCTTC 2820  
 TGATTACCCC CCATGTTGCA TATCAGGGTG CTCAAGGATT GGAGAGGAGA CAAAACCAGG 2880  
 25 AGCAGCACAG TGGGGACATC TCCCGTCTCA ACAGCCCCAG GCCTATGGGG GCTCTGGAAG 2940  
 GATGGGCCAG CTTGCAGGGG TTGGGGAGGG AGACATCCAG CTTGGGCTTT CCCCTTTGGA 3000  
 ATAAACCATT GGTCTGTC

Seq ID NO: 95 Protein sequence:  
 Protein Accession #: NP\_036233.1

30  
 35 1 MEAADASRSN 11 GSSPEARDAR 21 SPSGSPSGSLE 31 NGTKADGKDA 41 KTTNGHGGEA 51 AEGKSLGSAL 60  
 KPGEGRSALF AGNEWRRPII QFVESGDDKN SNYFSMDSME GKRSYPYAGLQ LGAAKKPPVT 120  
 FAEKGDVRKS IFSES RKPTV SIMEPGETRR NSYPRADTGL FSRSKSGSEE VLCDSCIGNK 180  
 QKAVKSCLVC QASFCEHLHLK PHLEGAAFRD HQLLEPIRDF EARKCPVHGK TMELFCQTDQ 240  
 TCICYLCMFQ EHKNHSTVTV EEAKAEKETE LSLQKEQLQL KIIIEIEDEAE KWQKEKDRIK 300  
 40 SFTTNEKAIL EQNFRDLVRD LEKQKEEVRA ALEQREQDAV DQVKVIMDAL DERAKVLHED 360  
 KQTRQLHSI SDSVLFLOEF GALMSNYSLP PPLPTYHVLL EGEGLGQSLG NFKDDLINVC 420  
 MRHVEKMKCA DLSRNFIERN HMENGGDHRY VNNYTNFSFGG EWSAPDTMKR YSMYLT PKGG 480  
 VRTSYQPSSP GRFTKETTKQ NFNNLYGTGK NYTSRVWEYS SSIQNSDNDL PVVQSSSFS 540  
 LKGYPSLMRS QSPKAQPQTW KSGKQTM LSH YRPFYVNKGN GIGSNEAP

45 Seq ID NO: 96 DNA sequence  
 Nucleic Acid Accession #: NM\_080668.1  
 Coding sequence: 83-841

50 1 GGCACGAGGG 11 CAGCGAGTGG 21 CCTTCCCGGT 31 TGGCGCGCGC 41 CCGGGGCGGC 51 GCGCTGGAG 60  
 GAGCTCGAGA CGGAGCCTAG TTATGTCTGG GAGGCGAACG CGGTCCGGAG GAGCCGCTCA 120  
 GCGCTCCGGG CCAAGGGCCC CATCTCCTAG TAAGCCTCTG CGGAGGTCCC AGCGGAAATC 180  
 55 AGGCTCTGAA CTCCCGAGCA TCCTCCCTGA AATCTGGCCG AAGACACCCA GTGCGGCTGC 240  
 AGTCAGAAAG CCCATCGTCT TAAAGAGGAT CGTGCCCAT GCTGTAGAGG TCCCAGCTGT 300  
 CCAATCACCT CGCAGGAGCC CTAGGATTTT CTTTCTCTTG GAGAAAGAAA ACGAGCCCCC 360  
 TGGCAGGGAG CTTACTAAGG AGGACCTTTT CAAGACACAC AGCGTCCCTG CCACCCCCAC 420  
 CAGCACTCCT GTGCCGAACC CTGAGGCCGA GTCCAGCTCC AAGGAAGGAG AGCTGGACGC 480  
 60 CAGAGACTTG GAAATGTCTA AGAAAGTCAG GCGTCTCTAC AGCCGGCTGG AGACCCTGGG 540  
 CTCTGCCTCT ACCTCCACCC CAGGCCGCGG GTCTGCTTT GGCTTCGAGG GGCTGCTGGG 600  
 GGCAGAAGAC TTGTCCGGAG TCTCGCCAGT GGTGTGCTCC AAACCTACCG AGGTCCCCAG 660  
 GGTTTGTGCA AAGCCCTGGG CCCCAGACAT GACTCTCCCT GGAATCTCCC CACCACCCGA 720  
 GAAACAGAAA CGTAAGAAGA AGAAAATGCC AGAGATCTTG AAAACGGAGC TGGATGAGTG 780  
 65 GGCTGCGGCC ATGAATGCCG AGTTTGAAGC TGCTGAGCAG TTTGATCTCC TGGTTGAATG 840  
 AGATGCAGTG GGGGGTGAC CTGGCCAGAC TCTCCCTCCT GTCTGTACA TAGCCACCTC 900  
 CCTGTGGAGA GGACACTTAG GGTCCCTCC CCGTCTCTTG TTACCTGTGT GTGTGCTGGT 960  
 GCTGCGCATG AGGACTGTCT GCCTTTGAGG GCTTGGGCAG CAGCGGCAGC CATCTTGTT 1020  
 TTAGGAAATG GGGCCGCTG GCCCAGCCAC TCACTGGTGT CCTGTCTCTT GTCGTCTGT 1080  
 70 CCTTCCTATC TCCCCAAGT ACCATAGCCA GTTTCAGAT GGGCCACAGA CTGGGGAGGA 1140  
 GAATCAGTGG CCCAGCCAGA AGTTAAAGGG CTGAGGGTTG AGGTGAGAGG CACCTCTGCT 1200  
 CTTGTTGGGA GGGGTGGCTG CTTGGAAATA GGCCAGGGG CTCTGCCAGC CTCGGCCTCT 1260  
 CCCTCCTGAG TTGCCTTCTG TTGGTGGCTT TCTTCTTGAA CCCACCTGTG TAAAGAGGTT 1320  
 75 TTCAGTTCCG TGGGTTTCCC CTTTGATTCT GTAAATAGTC CCAGAGAGAA TTCGTGGGCT 1380  
 GAGGGCAATT CTGTCTTGA GGAAGAAGCT GGACATTCAG CCTGTGGAGT CTGAGTTTGT 1440  
 AAGGATGTAG GGAGCCTTAG TTGGGTCTCA GACCATAAGT GTGTACTACA CAGAAGCTGT 1500  
 GTTTTCTAGT TCTGGTCTGC TGTGAGATG TTTGGTAAAT GCCAGGTTGA TAGGGCGCTG 1560  
 GCTGCTTGA GCAAAGGGTG CATTTCAGGG TGTGGCCACC AGGTGCTGTG AGTTTCTGTG 1620  
 GCTCATGGCC TGTGGGCTGG TCCCTTGCAC AGGGCCACG CTGGAGTCTT ACCACTCTGC 1680  
 80 TGCAGGGGTG GAAGTGGCC CCTCTGTCA CCCATACCCA TTTCTTACAA AATAAGTTAC 1740  
 ACCGAGTCTA CTTGGCCCTA GAAGAGAAAG TTGAAGAGTC CCAGACCTAC TAGCATTTTG 1800  
 CAACTATGCT TGTAAGTCC TCGGAAAGTT TCCTCGCGTA CCAGACAGCG GCGGGGGCTG 1860  
 ATAGCAATTT TAGTTTTGG CCTCCCTATC CTCTCACATG AGAACACTGC CTGGATGCAT 1920  
 CTCATGATCT CTGGAGAATT TCCCCATCTT TCTCTCTTT CCATCGTGTG GATTCAATAG 1980  
 85 TTTGGATTTG AAGGCTGCC TGCCCCGAC TCTCCTGCC CACCCCTGGC CATTGTACCT 2040  
 TTTGATGTTT AGAAGTTCGT GGAAGTAGAC GTCAGGTTGT GCAGAGGAGC TGGTGGATAA 2100  
 CAGAGAATGC CAGGGAAGAT GAGTGCTGGG TCAGGGTACT TGGATGAAAC GGTGCAGGCC 2160  
 AGGCGGGCCC TAATAAAACC CTCTGCCAGG TCTGGGAGTC CCAGGCCATC TGCTCAACGC 2220